

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : A01K 67/027, C12N 9/16		A1	(11) International Publication Number: WO 00/64247 (43) International Publication Date: 2 November 2000 (02.11.00)
(21) International Application Number: PCT/CA00/00430 (22) International Filing Date: 20 April 2000 (20.04.00)		(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).	
(30) Priority Data: 60/130,508 23 April 1999 (23.04.99) US		(71) Applicant (for all designated States except US): UNIVERSITY OF GUELPH [CA/CA]; College of Biological Science, Guelph, Ontario N1G 2W1 (CA).	
(72) Inventors; and (75) Inventors/Applicants (for US only): FORSBERG, Cecil, W. [CA/CA]; 44 Hands Drive, Guelph, Ontario N1G 2W1 (CA). GOLOVAN, Serguei [CA/CA]; 89-252 Stone Road, Guelph, Ontario N1G 2V7 (CA). PHILLIPS, John, P. [US/CA]; General Delivery, Arkell, Ontario N0B 1C0 (CA).		Published With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.	
(74) Agents: CHARI, Santosh, K. et al; Orange & Chari, 4900-55 King Street West, P.O. Box 190, Toronto Dominion Bank Tower, Toronto-Dominion Centre, Toronto, Ontario MSK 1H6 (CA).			

(54) Title: TRANSGENIC ANIMALS EXPRESSING SALIVARY PROTEINS

(57) Abstract

The invention provides a transgenic animal having within its genome a transgene construct for gastrointestinal tract specific expression of a protein. In a preferred embodiment, the protein is a phytase or a homologue thereof. Such proteins may be heterologous and may be specifically expressed in the salivary gland of the animal by operably linking the nucleic acid sequence encoding the protein with regulatory sequence including a salivary gland protein promoter/enhancer. Also provided are methods of expressing and producing proteins using such nucleic acid constructs. Further, antibodies specific to such proteins and immunological diagnostic kits are also provided.

BEST AVAILABLE COPY

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

TRANSGENIC ANIMALS EXPRESSING SALIVARY PROTEINS

FIELD OF THE INVENTION

5 The present invention relates to transgenic animals and, more specifically, to animals genetically modified to express a desired protein.

BACKGROUND OF THE INVENTION

10 Phosphorus is an essential element for the growth of all organisms. In livestock production, phosphorus deficiency has been described as the most prevalent mineral deficiency throughout the world and feed must often be supplemented with inorganic phosphorus in order to obtain desired growth performance of monogastric animals (e.g. pigs, poultry etc.).

15 Phytic acid, or phytate, (*myo*-inositol 1,2,3,4,5,6-hexakis dihydrogen phosphate) is a major storage form of phosphorus in cereals and legumes, representing 18% to 88% of the total phosphorus content (Reddy *et al.* 1982). The enzyme phytase (*myo*-inositol hexakisphosphate phosphohydrolase) belongs to the group of phosphoric monoester hydrolases: it catalyzes the hydrolysis of phytate (*myo*-inositol hexakis phosphate) to inorganic monophosphate and lower phosphoric esters of *myo*-inositol or, in some cases, free 20 *myo*-inositol. Phytases are classified either as 3-phytases or 6-phytases based on the first phosphate group attacked by the enzyme. 3-phytase is typical for microorganisms and 6-phytase for plants (Cosgrove, 1980).

25 Phytase is either absent or present at a very low levels in monogastric animals (Bitar and Reinhold 1972; Iqbal *et al.* 1994). Consequently, dietary phytate is not digested or absorbed from the small intestine and instead is concentrated in fecal material, thereby contributing to phosphorus pollution in areas of intensive livestock production. Runoff from animal farms leads to contamination of rivers and streams. Such runoff has resulted in rapid drops in the oxygen concentration in rivers and streams due to excessive algal growth in water, which, in turn, has led to an increase in the mortality rate of fish and existing flora and fauna. This is becoming a global problem as pig and poultry production is increased (Miner 30 1999; Mallin 2000). Furthermore, phytic acid is viewed as an anti-nutritional factor because it interacts with essential dietary minerals and proteins limiting the nutritional values of cereals and legumes in man and animals (Harland and Morris 1995).

For the above reasons, various attempts have been made to enable animals to utilize available phytate in feed. Such attempts have included production of low phytate plants (Abelson 1999), addition of phytase to the animal feed (Simons *et al.* 1990) (Stahl *et al.* 1999) or transformation of the fodder plants to produce the required phytase (Pen *et al.* 1993, Verwoerd *et al.* 1995). A combination of these options, the feeding of phytase to poultry receiving low phytate corn has also been tested (Huff *et al.* 1998). However, these solutions increase the cost of animal production. Also because phytase is an enzyme, it is susceptible to inactivation by heat and moisture and is generally unstable at the high temperatures used for feed pelleting.

10 The primary phytase used for supplementing animal feeds is from *Aspergillus* sp.; however, phytases are produced by a large number of plants and microorganisms (Wodzinski and Ullah 1996) (Dvorakova 1998). A phytase produced by *Escherichia coli* has been reported to exhibit the highest activity of those reported (Wodzinski and Ullah 1996). This phytase from *E. coli* was initially cloned as an acid phosphatase gene that was designated 15 *APPA* (Dassa *et al.* 1990). Greiner *et al.* (1991; 1993) purified phytase from *E. coli* and reported that some of the kinetic properties of the acid phosphatase activity of the native phytase of *E. coli* were similar to those of the *APPA*-encoded acid phosphatase. However, the authors did not clone the phytase gene to prove that it was identical to *APPA* gene. We have subsequently cloned, overexpressed and characterized *APPA* gene, and shown that the 20 *E. coli* gene *APPA* codes for a bifunctional enzyme exhibiting both phytase and acid phosphatase activities (Golovan *et al.* 2000). Phytases exhibit phosphatase activity, however the relative activities differ widely among enzymes (Wodzinski and Ullah 1996).

Therefore, there is a need for an improved method of allowing access by animals to phytase so as to enable efficient phytate metabolism and, thereby reducing phosphate pollution.

In the field of protein production using recombinant methods, one of the associated problems relates to the lack of required glycosylation. Therefore, a method of producing such glycoproteins is also needed.

30 SUMMARY OF THE INVENTION

In one embodiment, the invention provides a transgenic non-human animal that carries in the genome of its somatic and/or germ cells a nucleic acid sequence including a heterologous transgene construct, the construct including a transgene encoding a protein, the

transgene being operably linked to a first regulatory sequence for salivary gland specific expression of the protein.

5 In another embodiment, the invention provides a transgenic non-human animal that carries in the genome of its somatic and/or germ cells a nucleic acid sequence including a heterologous transgene construct, the construct including a transgene encoding phytase or a homologue thereof.

In yet another embodiment, the invention provides a method of expressing a protein, the method comprising the steps of:

10 a) introducing a transgene construct into a non-human animal embryo such that a non-human transgenic animal that develops from the embryo has a genome that comprises the transgene construct, wherein the transgene construct comprises:

- i) a transgene encoding the protein, and
- ii) at least one regulatory sequence for gastrointestinal tract specific expression of the protein,

15 b) transferring the embryo to a foster female; and,
c) developing the embryo into the transgenic animal

wherein the transgene is produced in the gastrointestinal tract of the animal.

20 In a further embodiment, the invention provides a transgenic animal adapted for expressing a protein according to the above method. The invention also provides for the progeny of such animal.

In another embodiment, the invention provides a process for producing a protein comprising the steps of:

25 a) obtaining saliva containing the protein from a non-human transgenic animal, the animal containing within its genome a transgene construct, wherein the transgene construct comprises:

- i) a transgene encoding the protein, and
- ii) at least one regulatory sequence for salivary gland specific expression of the protein, and

extracting the protein from the saliva.

30 In a further embodiment, the invention provides a method for expressing a phytase or a homologue thereof in a non-human animal, the method comprising:

a) constructing a nucleic acid sequence including a transgene construct comprising:
i) a transgene encoding the phytase or a homologue thereof, and

ii) at least one regulatory sequence for gastrointestinal tract specific expression of the protein, and

b) transfecting the animal with the nucleic acid sequence;

whereby the animal carries within the genome of its somatic and/or germ cells the transgene 5 construct and wherein the animal expresses the phytase or a homologue thereof in its gastrointestinal tract.

In another embodiment the invention provides a nucleic acid molecule comprising a nucleic acid sequence including a gene encoding a protein, the gene being operably linked to at least one regulatory sequence for gastrointestinal tract specific expression of the protein.

10 In another embodiment the invention provides an antibody specific to the protein expressed by the above nucleic acid sequence and a test kit for immunologically detecting such protein. The invention also provides for hybridomas secreting such antibodies.

15 In another embodiment the invention provides cells that are transfected with the above nucleic acid sequence.

20 In another embodiment, the invention provides a method for producing a protein molecule comprising a glycosylated protein secreted in the saliva that exhibits a novel physiological activity. One example of such an activity is phytase.

BRIEF DESCRIPTION OF THE DRAWINGS

25 These and other features of the preferred embodiments of the invention will become more apparent in the following detailed description in which reference is made to the appended drawings wherein:

Figure 1 is a schematic diagram representing a method for producing the gene construct of the present invention containing the inducible proline-rich protein (PRP) promoter/enhancer. More specifically, Figure 1 is a schematic diagram illustrating the steps in the construction of the transgenes R15/APPA+intron and R15/APPA used for the 30 generation of transgenic mice.

Figure 2 is a schematic diagram representing a method for producing the gene construct of the present invention containing the SV40 promoter. More specifically, Figure 2 is a schematic diagram illustrating the steps in construction of the plasmid containing the transgene SV40/APPA+intron that was introduced by transfection into mammalian cell lines.

Figure 3 is a schematic diagram representing a method for producing the gene construct of the present invention containing the constitutive parotid secretory protein (PSP) promoter/enhancer. More specifically, Figure 3 is a schematic diagram illustrating the steps

in construction of the transgenes *Lama2/APP*A that codes for the native *AppA* phytase and the *Lama2/PSP/APP*A that codes for the *AppA* phytase with the PSP signal peptide sequence.

Figure 4 is a schematic diagram of the *Lama2-APP*A plasmid containing the *APP*A transgene.

5 Figure 5 illustrates the nucleic acid sequence of the *Lama2/APP*A plasmid containing the *E. coli APPA* gene (SEQ ID NO: 1).

Figure 6 illustrates the PCR results for transformed mice. More specifically, figure 6 is a picture of an agarose gel illustrating *APP*A PCR products from genomic tail DNA of third generation offspring from the transgenic female founder mouse 3-1 generated using the 10 *Xba*1 and *Not*1 fragment of the *Lama2/APP*A construct. A second generation phytase gene positive male was crossed with each of two phytase positive transgenic females 9f and 11f (Table 3). From litter 18m x 9f offspring 3, 4, 5 & 6 are PCR positive and from litter 18m x 15 11f offspring 2 and 3 are PCR positive. Std is the oligonucleotide standard and the numbers on the left are the bp sizes of the standard. Lane C is a negative control reaction mixture that lacks a DNA template and *appA* is a positive control containing an amplified segment of the phytase gene. The primers used were *APP*A-UP2 and *APP*A-KPN.

Figure 7 illustrates the PCR results for transformed founder pigs. More specifically, Figure 7 is a picture of an agarose gel illustrating phytase gene PCR products and β -globin PCR products from genomic tail DNA of five founder piglets from litter 167. Std is a 1 kb ladder. Lane 2 using the phytase primer set is positive for the phytase gene, and all of the samples were positive for the β -globin gene. Lane C is a negative control not containing template DNA. The phytase transgene primer set included *APP*A-UP2 and *APP*A-KPN gave an expected fragment size of 750 bp. The primer set for the β -globin gene included PIG-BGF and PIG-BRG gives an expected fragment size of 207 bp.

25 Figure 8 illustrates the PCR results for transgene rearrangement tests. More specifically, Figure 8 is a picture of an agarose gel showing the PCR products of four separate primer sets used to amplify different segments of the transgene introduced into pig 167-02. The Std contained a kilobase DNA ladder. The primers used included lane 1, *APP*A-UP2 and *APP*A-KPN (750 bp); lane 2, *APP*A-MATURE and *APP*A-KPN (1235 bp); lane 3 *APP*A MATURE and *APP*A-DOWN2 (608 bp); lane 4, PIG-BGF and PIG-BGR (207 bp). 30 lane 5, a negative control without DNA template added; lane 6, the *appA* gene & primers *APP*A-UP2 and *APP*A-KPN. The numbers on the left indicate the sizes of the bands in the standard. No PCR products were detected in the absence of either DNA template or primers.

Figure 9 illustrates weight and salivary phytase activity of the transgenic boar 167-02 and average weight of the pen-mates at intervals during growth. Symbols: Weight of 167-02, ●; Average weight ± SD of four penmates, ▲; phytase activity of 167-02, ■; Phytase specific activity, □. Arrows indicate sampling for fecal phosphorus concentration.

5 Figure 10 illustrates weight and salivary phytase activity of the transgenic boar 282-02 and average weight of the pen-mates at intervals during growth. Symbols: Weight of 282-02, ●; Average weight ± SD of five penmates, ▲; phytase activity of 282-02, ■; Phytase specific activity, □. Arrows indicate sampling for fecal phosphorus concentration.

10 Figure 11 illustrates weight and salivary phytase activity of the transgenic boar 282-04 and average weight of the pen-mates at intervals during growth. Symbols: Weight of 282-04, ●; Average weight ± SD of five penmates, ▲; phytase activity of 282-04, ■; Phytase specific activity, □. Arrows indicate sampling for fecal phosphorus concentration.

15 Figure 12 illustrates weight and salivary phytase activity of the transgenic boar 405-02 and average weight of the pen-mates at intervals during growth. Symbols: Weight of 405-02, ●; Average weight ± SD of four penmates, ▲; phytase activity of 405-02, ■; Phytase specific activity, □. Arrows indicate sampling for fecal phosphorus concentration.

20 Figure 13 illustrates weight and salivary phytase activity of the transgenic boar 421-06 and average weight of the pen-mates at intervals during growth. Symbols: Weight of 421-06, ●; Average weight ± SD of four penmates, ▲; phytase activity of 421-06, ■; Phytase specific activity, □. Arrows indicate sampling for fecal phosphorus concentration.

25 Figure 14 illustrates the PCR results of first generation pigs. More specifically, Figure 14 is a picture of an agarose gel showing the PCR analysis of eight litter 154 piglets. The phytase transgenic boar 167-02 was used to breed a non-transgenic female. Std, 100 bp ladder, numbers on left are the sizes of the fragments in each band in bp; lane 167-02, DNA from boar 167-02 1, DNA from 167-02; lane C, is a lane without added DNA; lanes 1-8, are amplified DNA inserts from each of the offspring piglets of the litter. Phytase primers were Lama-UP and APPA-DOWN4. β -globin primers were PIG-BGF and PIG-BGR.

30 Figure 15 illustrates a sodium dodecylsulfate gel stained with silver demonstrating the sizes of the *E. coli* produced APPA phytase and the APPA phytase produced by the pig and a demonstration that the pig phytase is glycosylated. More specifically, Figure 15 is a picture of a sodium dodecylsulfate polyacrylamide gel electrophoresis (SDS-PAGE) profile of the purified AppA phytase produced in *E. coli* and the purified pig salivary phytase stained directly with silver (A) and a transfer from a similar SDS-PAGE gel transferred to

nitrocellulose and stained for glyoproteins (B). Creatinase is not glycosylated while transferring is glycosylated. The numbers on the left are the masses in of the molecular mass standards (Std) expressed in kDa.

Figure 15B is a picture of Western blot of the untreated pig AppA phytase and the 5 same phytase after treatment with a combination of three deglycosylating enzymes. Lane 1, Purified AppA phytase produced in *E. coli* (untreated); lane 2, purified pig phytase (untreated); lane 3, purified pig phytase treated with the combination of deglycosylating enzymes including N-glycosidase F, O-glycosidase and neuraminidase.

Figure 16 illustrates a Western blot of the pig phytase and the *E. coli* produced APPA 10 phytase using monoclonal antibodies directed to the APPA phytase documenting that they have homologous epitopes. More specifically, Figure 6 is a Western blot of the AppA phytase from pig saliva after various purification steps and of purified phytase produced in *E. coli*. A monoclonal antibody prepared against the *E. coli* phytase was used as the primary antibody for detection. Lane 1, saliva from non-transgenic pig 164-04; lane 2, saliva from 15 transgenic pig 167-02; Lane 3, saliva fraction not bound to DEAE-Sepharose; lane 4, salivary phytase bound to DEAE-Sepharose and released with an NaCl gradient; lane 5, salivary phytase further purified by Chromatofocusing with a pH gradient of 4 to 7; lane 6, phytase purified from *E. coli*. The numbers on the left are the masses of molecular mass standards (not shown) expressed in kDa.

Figure 17 illustrates an SDS-Page of the *E. coli* APPA phytase, saliva samples from 20 phytase negative and positive pigs and mice and a corresponding Western blot documenting that phytases from all three sources have homologous antigenic epitopes, but the animal phytases are larger than that produced in *E. coli*. More specifically, Figure 6 is a SDS-PAGE profile of the purified *E. coli* produced AppA phytase and the AppA phytases produced by 25 pigs and mice stained with silver (A) and a Western blot of an identical set of protein samples (B). A polyclonal antibody prepared against the *E. coli* phytase was used as the primary antibody for detection. Lane 1, Purified AppA phytase produced in *E. coli*; lane 2, Saliva from a non-transgenic pig 164-01; lane 3, Saliva from a AppA producing transgenic pig 167-02; lane 4, Purified phytase from pig 167-02; lane 5, Saliva from a non-transgenic mouse; 30 lane 6, Saliva from a transgenic mouse containing R15/APPA transgene induced with isoproterenol; lane 7, Saliva from a transgenic mouse containing the Lama/APPA transgene; Std, Molecular mass markers. The numbers on the left are the masses of molecular mass standards (not shown) expressed in kDa.

Figure 18 illustrates the nucleic acid sequence of the known segment of the R15/APPA + intron plasmid including the vector sequences of pBLCAT3 (SEQ ID NO:2).

Figure 19 illustrates the nucleic acid sequence of the known segment of the R15/APPA + intron transgene construct used for the generation of transgenic mice (SEQ ID NO:3).

Figure 20 illustrates the nucleic acid sequence of the known segment of the R15/APPA plasmid including the vector sequences of pBLCAT3 (SEQ ID NO:4).

Figure 21 illustrates the nucleic acid sequence of the known segment of the R15/APPA transgene construct used for the generation of transgenic mice (SEQ ID NO:5).

Figure 22 illustrates the nucleic acid sequence of the SV40/APPA + intron plasmid (SEQ ID NO:6).

Figure 23 illustrates the nucleic acid sequence of the Lama2/APPA transgene construct used for the generation of transgenic mice and transgenic pigs (SEQ ID NO: 7).

15 DESCRIPTION OF THE PREFERRED EMBODIMENTS

In the following description, a number of recombinant DNA technology terms are used. The following definitions have been provided in order to enable a clearer understanding of the specification and appended claims:

"Promoter" - a DNA sequence generally described as the 5' region of a gene and located proximal to the start codon. The transcription of an adjacent gene is initiated at the promoter region. If a promoter is an inducible promoter then the rate of transcription increases in response to an inducing agent. A constitutive promoter is one that initiates transcription of an adjacent gene without additional regulation.

"Operably Linked" - a nucleic acid sequence is "operably linked" when placed into a functional relationship with another nucleic acid sequence. For instance, a promoter or enhancer is "operably linked" to a coding sequence if the promoter causes the transcription of the sequence. Generally, operably linked means that the linked nucleic acid sequences are contiguous and, where it is necessary to join two protein coding regions, contiguous and in one reading frame.

"Phytase" - any protein that liberates phosphate from myo-inositolhexakis-phosphate or other inositol phosphates. Its catalytic capability may be limited to phytic acid or one of its salts, or it may show less specificity and hydrolyze a variety of phosphorylated compounds.

"Gene" - a DNA sequence that contains a template for an RNA polymerase and contains information needed for expressing a polypeptide or protein.

"Polynucleotide Molecule" - a polydeoxyribonucleic (DNA) acid molecule or a polyribonucleic acid (RNA) molecule.

5 "Expression" - the process by which a polypeptide is produced from a structural gene.

"Cloning vehicle" - is a plasmid or phage DNA or other DNA sequence which is capable of carrying genetic information into a host cell. A cloning vehicle is often characterized by one or more endonuclease recognition sites at which such DNA sequences may be cut in a determinable fashion without loss of an essential biological function of the 10 vehicle. A cloning vehicle is a DNA sequence into which a desired DNA may be spliced in order to bring about its cloning into the host cell.

"Vector" - is a term also used to refer to a cloning vehicle.

"Plasmid" - is a cloning vehicle generally comprising a circular DNA molecule that is maintained and replicates autonomously in at least one host cell.

15 "Expression vehicle" - a vehicle or vector similar to a cloning vehicle but which supports expression of a gene that has been cloned into it, after transformation of a host. The cloned gene is usually placed under the control of (i.e. is operably linked to) certain control sequences such as promoter sequences.

"Host" - a cell that is utilized as the recipient and carrier of recombinant material.

20 "Homologous" - refers to a nucleic acid molecule that originates from the same genus or species as the host.

"Heterologous" - refers to a nucleic acid molecule that originates from a different genus or species than that of the host.

"Glycoprotein" - refers to a peptide molecule that has undergone glycosylation.

25 "Glycosylation" - refers to the addition of carbohydrate groups to a amino acid residues of a peptide molecule.

In recent years, transgenic animals have been developed for many purposes (Pinkert *et al.* 1990) (Wall *et al.* 1997). One premise, therefore, for the present invention is that by providing a transgenic animal capable of expressing phytase, the problems discussed above 30 would be obviated. The options for heterologous phytase expression in animals include (i) salivary gland secretion of a phytase, (ii) pancreatic secretion of the enzyme into the small intestine along with the digestive enzymes, or (iii) secretion from the intestinal epithelial cells much like that of indigenous alkaline phosphatase and glycosidases (Low, 1989). The *E. coli* phytase would appear to be best suited for hydrolytic activity in the monogastric stomach

because the enzyme has a pH optimum in the range of 2.5 to 4.5 and it is resistant to pepsin, the predominant protease active in the stomach. The phytase has a periplasmic location in *E. coli* and has an N-terminal signal peptide sequence (Golovan et al., 1999) that seemed optimally adapted for secretion from the parotid gland. Phytase could be expressed in either the pancreas 5 for secretion into the small intestine or it could be expressed in the intestinal epithelial tissue and secreted into the intestinal milieu. However, if these choices of expression locations were chosen, it would be necessary to select an enzyme active at the more neutral pH of the small intestine and one which was more resistant to pancreatic enzymes including trypsin, chymotrypsin and elastase.

10 Factors of importance in terms of the expressed enzyme when selecting a phytase for expression in the gastrointestinal tract include a pH that is optimum for activity, high catalytic activity, broad substrate specificity, and protease resistance. If any of these properties, or indeed others, is not acceptable, there are now sophisticated molecular methods for modifying the properties of an enzyme. These include site directed mutagenesis, random 15 mutagenesis and various modifications of DNA shuffling (Harayama, 1998; Crameri et al., 1998).

Synthesis of phytase in the salivary gland and secretion in the saliva would, therefore, provide for early contact of the enzyme with phytic acid present in the feed and provide sufficient time for hydrolysis.

20 The salivary gland system of the pig consists of three pairs of glands, the parotid gland, which secretes through a duct on each cheek, and mandibular and submaxillary glands that have joint ducts that secrete beneath the front on the tongue. Saliva secreted in the pig via these ducts is discontinuous and is produced during consumption of solid foods, and can equal the weight of food consumed when water is limited during feed consumption (Corring, 1980; Arkhipovets, 25 1956). For example, the quantity of saliva produced by a 45 kg pig can vary from near zero when the pig receives a mainly liquid diet to 500 g when a dry diet is consumed without access to water. The salivary glands of the pig secrete amylase (Rozhkov and Galimov, 1990) and a variety of other salivary proteins and mucopolysaccharides.

To our knowledge no porcine genes coding for salivary proteins have been cloned.

30 However, genes coding for major proteins secreted by the rat and mouse have been cloned and characterized. A multigene family encoding a group of unique proteins high in proline, the so-called proline-rich proteins (PRPs) are produced when either mice or rats consume tannins or are injected with isoproterenol.

It would be advantageous to develop an animal that is transformed to express phytase, preferably in the salivary gland. In such case, the phytate naturally occurring in the animal feed can be utilized by the animal without any additives being used. This will decrease the cost of animal production, and furthermore, will avoid polluting the environment with phosphorus. Therefore, the present invention aims to overcome the deficiencies of the prior art relating to increasing phytate utilization and, particularly, to provide transgenic animals which express phytase.

In the production of heterologous proteins by means of recombinant methods, several hurdles have been faced. One such hurdle that is often faced is the lack of required post-translational modification of the expressed protein thereby resulting in a protein that is structurally and/or functionally, different from the desired molecule. Glycosylation is one such post-translational modification that is desired. However, such modification is generally found to occur in more complex mammalian systems. Therefore in one embodiment of the present invention there is provided a method of producing recombinant glycoproteins.

In one embodiment, the present invention provides an animal capable of inducible or constitutive salivary expression of a heterologous protein. To illustrate this, the mouse was chosen as the animal model and the gene constructs used for transformation were created using the rat proline-rich protein (PRP) promoter/enhancer (inducible promoter) and the mouse parotid secretory protein (PSP) promoter/enhancer (constitutive promoter). In this illustration, phytase was used for expression in saliva.

After finding that an inducible phytase could be expressed in the parotid gland of mice the expression of the phytase transgene under the control of the constitutive PSP promoter was then tested. Two mice transgenic for the PSP construct were produced under contract at the University of Alabama.

Following the testing of the mice described above, transgenic pigs were developed by introduction into the genome a phytase transgene consisting of a constitutive promoter driving the synthesis of a highly active phytase. The pigs so generated were found to excrete less phosphorus in their feces than non-transgenic pigs.

30 Expression in the Salivary Glands

Saliva is a clear colorless fluid secreted by major salivary glands (parotid, submandibular, sublingual and minor salivary) that lubricates and cleans the oral structure, as well as initiates the process of digestion. The parotid glands are two of six major glands associated with the production of saliva. The parotid gland is composed mainly of two cell

types: acinar and interglobular duct cells. The acinar cells, which represent 75 to 85% of the tissue, are the sites of secretory protein synthesis (Frandsen and Spurgeon 1992). Two very abundant proteins are produced by these cells: α -amylase (AMY-1) (2% of polyA RNA) (Madsen and Hjorth 1985), and parotid secretory protein (PSP) (10% of polyA RNA) (Shaw and Schibler 1986). Several constructs are now available which allow tissue-specific expression of a transgene in the salivary glands of mice.

The salivary secretion in pigs has not received the attention given to that of mice and humans. It was suggested that salivary secretion is discontinuous (less secreted between periods of meal consumption). Up to 500 g of saliva may be secreted by a 45 kg pig upon 10 consumption of 500 g of dry feed (Corring 1980). Wide variations were detected in both the flow rate and electrolytes in saliva between animals and even between samples taken from the same animal on separate days (Tryon and Bibby 1966). Very little is known about the composition of pig's saliva or salivary enzymes. Salivary amylase was detected, although the quantity was 250 000 times less than that of pancreatic amylase, and 100 times less than in 15 human saliva (Low 1989). There are no constructs known which would allow salivary gland-specific expression of transgene in pigs.

I) APPA Gene Under Control Of An Inducible Promoter

20 1) Construction of R15/APPA constructs (Inducible Promoter)

In this process, a plasmid is constructed by linking a promoter/enhancer for a saliva protein with the *APPA* gene, which codes for the bifunctional phytase, acid phosphatase. The *APPA* gene used in this construction was cloned from *E. coli* ATCC 33965 into pBR322. This is described above (Golovan et al., 2000).

25 Proteins, unusually high in proline, the so-called proline-rich proteins (PRPs), comprise about 70% of the total proteins in human saliva (Bennick 1982). Unlike the constitutive expression of the PRPs in humans, the salivary glands of mice, rats and hamster normally either do not express PRPs or express them in low levels. In the rat and mouse, PRP gene expression can be dramatically induced by diets high in tannins or by injection 30 with the β -agonist isoproterenol (Carlson 1993). After 6 to 10 days of daily isoproterenol injection the PRPs comprised about 70% of the total soluble protein in parotid gland extracts. PRP cDNA and PRP genes have been cloned and characterized from rats (Clements et al.

1985), mice (Ann and Carlson 1985), hamsters (Mehansho *et al.* 1987), and humans (Kim and Maeda 1986).

Transgenic mice were used to locate the cis-acting DNA elements that are essential for salivary-specific and inducible expression of the rat proline-rich protein gene, R15. It was 5 found that a parotid control region (-6 to -1.7 kb) upstream of the R15 promoter is capable of directing parotid-specific and isoproterenol-inducible expression of a heterologous promoter construct (Tu *et al.* 1993). The distal -10 to -6 kb region was shown to function as an enhancer, which can increase levels of expression more than 30-fold. The -6 to -1.7 kb region also seems to function as a locus control region (LCR), because it conferred copy 10 number-dependent and chromosomal position-independent expression of a reporter gene in 15 out of 15 independent transgenic mice (Tu, Lazowski, Ehlenfeldt, Wu, Lin, Kousvelari, and Ann 1993).

We obtained the R15-PRP promoter from Dr. D.K. Ann as a plasmid -10R15/CAT, which placed the chloramphenicol acetyltransferase gene (CAT) under control of the 15 inducible R15-PRP promoter. We decided to use the plasmid as a basis for transgene construction (Figure 1). Due to the absence of complete sequence information about the R15-PRP promoter (only 2 kbp out of 10 kbp was sequenced) we removed the R15-PRP promoter by Xho I digestion (Figure 1, step 1). Re-ligated plasmid was used as a template for PCR with CAT-ATG and CAT-TAA synthetic primers. The 4.3 kbp CAT_{PCR} fragment 20 had the initiation site of the CAT gene substituted with the optimal eukaryotic initiation sequence (Kozak 1987). The fragment was purified by agarose gel electrophoresis, re-ligated to itself and used to transform *E. coli* (Figure 1, step 2). The CAT_{PCR} plasmid was digested with Nco I and filled-in using T4 DNA polymerase to generate a blunt end. After that, the CAT_{PCR} fragment was digested with Eco47III and purified by agarose gel electrophoresis 25 (Figure 1, step 3). Three rare codons in the APPA gene were modified during the sub-cloning steps leading to the construction of the transgene. Specifically, the Ala₃ coding sequence was changed from GCG to GCC, the Pro₄₂₈ sequence was changed from CCG to CCC, and the Ala₄₂₉ sequence was changed from GCG to GCT. This modification was made in order to increase the possibility of transcription of the gene in eukaryotic cells. The APPA gene was 30 amplified by PCR using the previously cloned APPA gene from the pBR322/APPA plasmid with the synthetic primers APPA-DRA and APPA-SMA. The 1.3 kbp APPA_{PCR} fragment generated by PCR was digested with Dra I and Sma I and gel-purified (Figure 1, step 4). APPA_{PCR} and CAT_{PCR} fragments were blunt end ligated to produce CAT/APPA+intron

vector (Figure 1, step 5), which was introduced into a DH5 α strain of *E. coli*. The insert orientation was checked by restriction digest with Sal I and EcoR I. The transgene in CAT/APPA+intron was checked by sequencing both strands. To remove the SV40 small t intron the 2.3 kbp APPA/intron/polyA fragment was excised from a plasmid by Xho I and 5 EcoR I digestion (Figure 1, step 6a), gel purified and digested by Dra I (Figure 1, step 6b). The 1.5 kbp (APPA) and 0.2 kbp (polyA) fragments were gel-purified and linked together in three way ligation with CAT_{PCR} digested with Xho I and EcoR I (Figure 1, step 6c). The resulting plasmids CAT/APPA and CAT/APPA+intron were digested with Xho I, gel-purified and re-ligated with R15-PRP promoter digested with Xho I (Figure 1, step 7).

10 Because of the low efficiency of ligation the whole ligation mixture was used to transform *E.coli*, total plasmid DNA was prepared and run on the agarose gel. Plasmids which were larger than the original CAT/APPA (5.6 kbp) were eluted and re-transformed in *E.coli*. Plasmids with the R15-PRP insert (15 kbp) were identified by electrophoresing DNA from a single colony on an agarose gel. The correct orientation was identified by PCR with R15-

15 UP1 and APPA-DOWN2 synthetic primers. The plasmids R15/APPA and R15/APPA+intron were both digested with Hind III and Kpn I; transgenes were gel-purified and further purified using a Qiagen column (Figure 1, step 8).

Figure 18 illustrates the nucleic acid sequence for the plasmid containing the known segment of the R15/APPA + intron sequence including the vector sequences of pBLCAT3.

20 The sequence of this plasmid is designated as SEQ ID NO:2.

Figure 19 illustrates the nucleic acid sequence for the transgene construct containing the known segment of the R15/APPA + intron sequence used for the generation of transgenic mice. The sequence of this transgene is designated as SEQ ID NO:3.

25 Figure 20 illustrates the nucleic acid sequence for the plasmid containing the known segment of the R15/APPA sequence including the vector sequences of pBLCAT3. The sequence for this plasmid is designated as SEQ ID NO:4.

The pBLCAT3 sequence indicated above is present in the CAT/APPA of Figure 1 and in the CAT/APPA+intron of Figure 2. This sequence was part of the original -10R15/CAT and a portion of it was carried through in the construction process.

30 Figure 21 illustrates the nucleic acid sequence for the transgene construct containing the known segment of the R15/APPA sequence used for the generation of transgenic mice. The sequence of this transgene is designated as SEQ ID NO:5.

2) Expression of SV40/APPA+intron in Cell Culture

To produce an SV40/APPA plasmid for expression of *APPA* in cell culture, the SV40 promoter/enhancer was amplified by PCR from the pSV- β -galactosidase plasmid (Promega) using the synthetic primers SV-HIND and SV-XHO. The SV40 promoter/enhancer fragment 5 was digested with Xho I and Hind III, gel purified, and ligated into CAT/APPA digested with Xho I and Hind III (Figure 2).

Figure 22 illustrates nucleic acid sequence for the SV40/APPA + intron. The sequence for this plasmid is designated as SEQ ID NO:6.

We obtained a rat parotid acinar cell line PARC 5.8 (Quissell *et al.* 1998) that we 10 intended to use for transient expression of the phytase transgene. The purpose was to test the efficiency of different constructs for transgene expression and also to detect any deleterious effects of phytase expression before introduction into the animals. We tried transient expression of the *APPA* gene using R15/APPA and R15/APPA+intron constructs but because of low transfection efficiency and/or low expression levels, we were unable to detect either 15 phytase or β -galactosidase that we used as a control for transfection.

We exchanged the R15-PRP inducible promoter from the R15/APPA construct with the SV40 constitutive promoter-enhancer, which enables high level transient expression in different cell cultures. CHO, COS7 and HELA cell lines were screened for transient expression of the *APPA* phytase using the SV40 promoter/enhancer. All cell lines were 20 maintained on DMEM/F12 (Sigma) cell medium with 10 % (wt/vol) heat-inactivated fetal bovine serum at 37°C in 5% CO₂ and 95% air. Cells were grown to 70 % confluence before transfection. Two hours before transfection the medium was exchanged with fresh medium. Cells were transformed with 5 μ g of DNA per 60 mm culture plate (1:1 SV40/*APPA* and 25 SV40/ β -galactosidase) using the DNA-Calcium-Phosphate method of transfection (Gorman *et al.* 1983). After 6 hours of incubation the medium was removed and cells were subjected to glycerol shock for 3 min (Ausbel *et al.* 1992). Cells were washed with phosphate-buffered saline (PBS) and incubated in fresh medium under standard growth conditions. After 48 hours of incubation cell-free culture fluid was collected, the cells washed two times with PBS and lysed with 1ml of 1% (vol/vol) NP-40, 1mM disodium EDTA in Hanks balanced salts 30 (HBSS) for 1 hour at 4°C. The phytase assay was performed in a final volume of 100 μ l of 0.1 M sodium acetate/acetic acid buffer (pH 4.5) using sodium phytate (4 mM) as a substrate at 37°C. After 6 hours of incubation the reaction was stopped with 67 μ l ammonium molybdate/ammonium vanadate/nitric acid mixture and the concentration of liberated

inorganic phosphate determined at 405 nm (Engelen *et al.* 1994). One unit (U) of enzyme activity was the amount of the enzyme releasing 1 μ mol inorganic phosphate per minute. The assay was performed in triplicate. As a control for endogenous phytase activity, non-transfected cell lines were used.

5 We did not detect endogenous phytase activity in non-transfected cell lines. Phytase activity was detected in all transfected cell lines, with COS7 cells expressing a total of 0.35 U of phytase in cell-free culture fluid (4 ml) and 0.0034 U in the cell fraction (1.1 ml) obtained from the same plate. The phytase activity produced by COS7 cells was 7 times higher than that of CHO and 35 times more than the HE LA cell line. More than 99% of activity was
10 located in cell-free culture fluid, which suggests that the expressed enzyme was exported out of the cell using the bacterial signal sequence. We were unable to detect expression of cytoplasmic β -galactosidase, which we wanted to use as a control for transfection efficiency.

3) Expression of R15-PRP/APPA in Transgenic Mice

15 Transgenic mice were generated using the constructs R15/APPA and R15/APPA+intron by Dr. C.A. Pinkert at the NICHD Transgenic Mouse Development Facility (NTMDF), University of Alabama at Birmingham, Alabama. The procedures followed in generating the mice have been standardized by the NTMDF and further information concerning this can be obtained at: <http://transgenics.bhs.uab.edu/page1.htm>, the
20 content of which is incorporated herein by reference. This procedure involved the microinjection technique for transfecting mice with the desired nucleic acid sequence. To summarize, the sequences are microinjected into mouse zygotes and the surviving eggs are implanted into pseudopregnant recipient mice. The recipient mice then give birth to the resulting founder transgenic mice. It will be appreciated that various other methods of
25 generating transgenic mice may be used in the present invention.

The R15/APPA transgene in mice was detected by PCR using the primers CAT-UP1 and APPA-DOWN2 that gives rise to a 700 bp fragment using the standard PCR conditions, except that the hybridization step was set at 51°C for 40 seconds and the polymerization step was at 72°C for one minute.

30 For the R15/APPA construct 8 PCR positive founder mice were obtained of which 4 were males and 4 were females. Three of the founders did not pass the transgene to progeny and were probably mosaics. For R15/APPA+intron 5 PCR positive founder mice were obtained, 3 were males and 2 were females, and one of them was found to be mosaic. At 10

to 12 weeks of age PRP production in the PCR positive progeny from different lines was induced for 10 days by daily intraperitoneal (ip) injection of 1mg isoproterenol dissolved in 100 μ l sterile saline. To serve as a control several PCR negative progeny were also induced. No significant differences in weight were noticed between PCR positive and PCR negative progeny at either the beginning or end of the induction period. Saliva was collected before induction and at the end of the 10 day induction period.

To collect saliva, mice were lightly anesthetized with a ketamine/xylazine mixture (ip injection of 50 mg ketamine and 5 mg xylazine per kg body weight diluted in water) and saliva flow was induced by injection with pilocarpine/isoproterenol (ip injection of 0.5 mg pilocarpine and 2 mg isoproterenol per kg body weight dissolved in saline) (Hu *et al.* 1992). Between 100-250 μ l of saliva was collected from each mouse over a 30 min period beginning 5 min after the pilocarpine/isoproterenol injection.

The saliva was collected from each mouse by holding it in one hand and withdrawing saliva from the corner of the mouth with a 20 μ l pipetter. Collected saliva was transferred to a cold Eppendorf microcentrifuge tube containing 2 μ l of 0.5 M EDTA (pH 8.0) and 4 μ l of 10 mg/ml protease inhibitor Pefabloc (Boehringer Mannheim) dissolved in water. The tubes with saliva were kept on ice until assays were conducted. Phytase activity in the saliva was assayed as described for the SV40/APPA expressed in cell culture.

Phytase expression was not detected in either un-induced or in induced PCR negative mice. For PCR positive mice, phytase expression was not detected in those that were un-induced. However, phytase expression was observed for PCR positive mice that were induced. The results of this study are summarized in Table 1.

Even though it was possible to distinguish saliva from induced PCR positive from that of PCR negative mice in a phytase assay by a characteristic yellow color, saliva from some of the negative mice, when assayed, produced cloudiness that was impossible to remove by centrifugation and that affected spectrophotometer readings. We did not notice any gender differences in expression, both males and females were found to produce phytase in saliva. In three lines (all R15/APPA+intron) no phytase expression or very low level of expression (0.03-0.95 U/ml) was detected, in 4 lines the level of expression ranged from 7 to 87 U/ml, and two lines (both R15/APPA) produced very high levels of phytase in saliva, 252 and 547 U/ml.

These experiments demonstrated that phytase can be expressed at a very high level in the salivary glands of mice, without detrimental effects on the animals. We also were able to

produce progeny with an inducible salivary phytase from animals expressing the inducible phytase thereby documenting inheritance of the trait, and showing that the reproductive capability of animals was not affected. When the F2 generation of mice were tested for salivary phytase the level of phytase production was preserved.

5 Founders containing the transgene without the intron gave offspring that produced significantly higher levels of phytase. The SV40 intron in the R15/APPA+intron construct seems to cause a lower level of expression, and in three lines (A1f, A20f and B0m) the level of phytase was barely detectable. The level of phytase expression in A2m line (R15/APPA+intron) was 6.2 times lower than that of the B0m-intron line (R15/APPA).

10 Preliminary experiments showed that when the enzyme was analyzed by PAGE its size was increased from 42 kDa to 60 kDa. It is likely modified by glycosylation, but stable and active.

II) APPA Gene Under Control Of A Constitutive Promoter

15

1) Construction of the Lama2/APPA Transgene (Constitutive Promoter)

20 The murine parotid secretory protein (PSP) is the most abundantly expressed protein in the parotid gland of mice (Madsen and Hjorth 1985). After an hour of pulse labeling, the mouse parotid gland incorporates 65 to 85% of ¹⁴C-leucine into this single protein (Owerbach and Hjorth 1980). It was estimated that PSP mRNA accumulates up to 50,000 molecules per cell and that from 3 to 5 molecules of PSP are produced for every molecule of amylase (Madsen and Hjorth 1985). Despite the predominance of the PSP in saliva its function is not well characterized.

25 The single-copy gene coding for PSP has been cloned and characterized. It has two alleles PSP^a (Shaw and Schibler 1986) and PSP^b (Owerbach and Hjorth 1980). The PSP^b allele is also expressed in the sublingual gland, but at 1/10 of the level found in the parotid gland. It was shown that 4.6 kbp of 5' flanking sequence of PSP^b is sufficient for salivary gland specific expression. The level of sublingual expression approached 100% of the PSP mRNA level, whereas the parotid expression did not exceed 1% (Mikkelsen *et al.* 1992),
30 which demonstrates that regulatory sequences for sublingual and parotid expression are not identical. The level of expression was also dependent on the site of integration. The same construct was used for expression of the C-terminal chain of the human blood coagulation factor VIII, FVIII. A high level of FVIII mRNA was detected in the sublingual gland and a low level in the parotid gland. The transgenic lines also secreted the FVIII light chain into

saliva at the level of about 10 units per salivation (about 0.05 ml of saliva) (Mikkelsen et al., 1992). Later the same group achieved a high level of parotid-specific expression that was similar or even exceeded that of the endogenous gene by using 11.4 kbp of 5' flanking sequences and 2.5 kbp of 3' flanking sequences (Larsen et al. 1994). The expression also 5 seems to be position-independent and copy-number-dependent that could indicate the presence of a LCR in these sequences.

Lama 2 is a portion of the PSP gene and comprises an 18 kbp construct that is expressed in transgenic mice at up to 56% of the endogenous PSP gene.

Because a large part of Lama 2 had not been sequenced, the construct was first 10 disassembled and subcloned into pBluescript KS(+) and after incorporation of the APPA gene, the Lama 2 was reassembled back (Figure 3). We used unique enzymes RsrII and SmaI to remove a 3.4 kbp fragment from Lama2, which was subcloned into the multiple cloning site (MCS) of pBluescript II KS(+) that was previously digested with KpnI and SmaI, using a KpnI-RsrII adapter (Figure 3, step 1).

15 KpnI* RsrII

TGGGAGGTCTG

CATGACCCTCCAGCCAG

That allowed us to preserve the RsrII (CG/GWCCG) site and destroy the KpnI site (GGTAC/C> GGTAC/T), which would otherwise interfere with future cloning. The 20 pKS/Lama construct was digested with ApaI and KpnI and used in a three-way ligation with the modified APPA (Figure 3, step 2). We designed two PSP/APPA constructs. One construct APPA-signal/APPA (Figure 3, steps 3a-7a) had the original bacterial signal sequence from the APPA protein having the following amino acid sequence:

25 Met-Lys-Ala-Ile-Leu-Ile-Pro-Phe-Leu-Ser-Leu-Leu-Ile-Pro-Leu-Thr-Pro-Gln-Ser-Ala-Phe-Ala

We also modified a sequence near the ATG codon to resemble the optimal 30 mammalian Kozak sequence (GCC GCC A/GCC ATG G) (Kozak 1987), but we did not mutagenize the +4 position because it would change Lys to Glu in the signal sequence with possible deleterious consequences for protein export. This optimized sequence was used in our previous construct R15/APPA and led to high levels of phytase production. We checked the APPA bacterial signal sequence using the PSORT computer neural network trained on eukaryotic signal sequences and further described at <http://psort.nibb.ac.jp:8800/> (Nakai and

Kanehisa 1992). The APPA bacterial signal sequence was recognized as an efficient leader peptide and the cleavage site was correctly predicted. PSORT also predicted that there is a high probability that phytase would be exported correctly outside of the cell. There were also publications showing that some bacterial signal sequences might function efficiently in 5 mammalian cells (Williamson *et al.* 1994) (Hall *et al.* 1990). Our experiments using cell culture demonstrated that the APPA signal was correctly processed with export of phytase outside of the cell.

Experiments using cell culture cannot predict the direction of export and if phytase were exported into blood vessels instead of salivary ducts that could lead to deleterious 10 effects. That is why we also designed a second construct PSP-signal/APPA (Figure 3, steps 3b-7b) that would preserve the original PSP signal amino acid sequence:

Met-Phe-Gln-Leu-Gly-Ser-Leu-Val-Val-Leu-Cys-Gly-Leu-Leu-Ile-Gly-Asn-Ser-Glu-Ser

15 This leader peptide was also efficiently recognized by PSORT with the correct cleavage site (Nakai and Kanehisa 1992). In this construct we also preserved the original PSP sequences near the ATG start codons, which may not be optimal, but could be important in regulation of gene expression. The APPA gene for both constructs was amplified by PCR using as the template our previous transgenic construct R15/APPA that possessed the optimal 20 Kozak sequence and the modified codons for residues Ala3, Pro428 and Ala429 as described earlier. For the APPA signal/APPA construct two synthetic primers were used which introduced a Clal site near the ATG codon (APPA-CLA) and a Kpn1 site near the TAA stop codon (APPA-KPN). The APPA_{PCR1} product was digested with Clal and Kpn1. The Clal site was also introduced into Lama 2 using pKS/Lama 2 as template for PCR. LAMA-UP 25 primer was located upstream of Apal site and the LAMA-CLA primer introduced the Clal site near ATG codon (Figure 3, step 3a). Lama_{PCR1} product was digested with Clal and Apal (Figure 3, step 4a). pKS/Lama (Apal-Kpn1), Lama_{PCR1} (Apal- Clal) and APPA_{PCR1} (Clal-Kpn1) were combined together in a three-way ligation reaction (Figure 3, step 5a). The recovered pKS/Lama/APPA plasmid was digested with RsrII, SmaI and inserted back 30 into Lama2 (Figure 3, step 6a).

For the PSPsignal/ APPA construct, the synthetic APPA -KPN primer was used with the synthetic APPA -MATURE primer, which produced phytase without a signal sequence. The APPA_{PCR2} product was blunt-ended using T4 DNA polymerase and digested with Kpn1. The PSP signal sequence was produced using the LAMA-UP and LAMA -SIGNAL primer

(Figure 3, step 3b). The $\text{Lama}_{\text{PCR}}2$ was blunt-ended using T4 DNA polymerase and digested with ApaI (Figure 3, step 4b). pKS/Lama (Apa1-Kpn1), $\text{Lama}_{\text{PCR}}2$ (Apa1-blunt) and APPA $_{\text{PCR}}2$ (blunt-Kpn1) were combined together in a three-way ligation reaction (Figure 3, step 5b). The recovered pKS/Lama/APPA plasmid was digested with RsrII, SmaI and

5 inserted back into Lama2 (Figure 3, step 6b).

Even though both constructs were successfully produced we decided to use Lama2/APPA signal/APPA for the generation of transgenic mice, because we have results from our previous transgenic constructs R15/APPA and R15/APPA+intron which demonstrated that phytase with optimized Kozak sequence and the APPA signal peptide was 10 synthesized at a high level in salivary glands after induction and was efficiently exported into the salivary duct. The Lama2/APPA vector was digested with XhoI and NotI, and the transgene was gel-purified and further purified using a Qiagen column (Figure 3, step 7a).

2) Sequence of the Lama2/APPA Construct

A large segment of the Lama2 construct (Laursen and Hjorth 1997) used for 15 construction of the Lama2-APPA transgene had not been reported in GenBank prior to our research. To ensure that we could more clearly describe the transgene construct, and furthermore to avoid the introduction of deleterious DNA sequences from the mouse into the pig in the process of generating transgenic pigs, we sequenced the Lama2-APPA plasmid on both strands. Figure 4 illustrates schematically the structure of the Lama2-APPA plasmid.

20 Figure 5 illustrates the nucleic acid sequence (SEQ ID NO:1) of such plasmid. The full transgene sequence was reconstructed from overlapping DNA sequences using the Contig Assembly Program (CAP) (<http://hercules.tigem.it/ASSEMBLY/assemble.html>) developed by Huang (1996; 1999) and then inspected manually for sequencing errors. The transgene sequence was checked for the presence of interspersed repetitive elements using the computer 25 program RepeatMasker (Smith and Green, RepeatMasker at <http://ftp.genome.washington.edu/cgi-bin/RepeatMasker>). It was found that 26 % of the transgene sequence was composed of repetitive elements (Table 2). However, such repetitive elements are widely present in all mammalian genomes. For example, up to 50% of the human genome is derived from repetitive elements (Smit 1996; Kazazian 1998).

30 Figure 23 illustrates the nucleic acid sequence (SEQ ID NO:7) of the Lama2/APPA transgene construct.

The Lama2 high level expression cassette (Laursen and Hjorth 1997) contains the enhancer region and the promoter of the *Psp* gene in the parotid gland. High expression was

shown to be dependent on regulatory elements between -11.5 kb and -6.5 kb and/or between +8.3 kb and +10.9 kb. Svendsen et al. (1998a) showed that a 1.5 kb sequence between -3.1 kb and -4.6 kb had properties of a parotid and sublingual specific enhancer and was designated as the PSP proximal enhancer. Furthermore, they showed that transgenes 5 containing the PSP promoter and 5' flanking region located between -3.6 kb and -4.3 kb contained sequence information necessary to direct salivary gland specific expression.

Screening the transgene with RepeatMasker did not reveal the presence of any full-length active autonomous elements. The repeats present were extensively modified by insertions and deletions. The *blastx* program was also used to compare the transgene 10 sequence translated in all reading frames against the National Center for Biotechnology Information (NCBI) protein sequence database (<http://www.ncbi.nlm.nih.gov/BLAST/>) (Altschul et al. 1990; Gish and States 1993; Terada and Nakanuma 1993). A region of DNA from 861 to 2180 was found that might code for parts of a protein with limited homology (38-58% identities) to the C-terminus of several human and mouse reverse transcriptases. 15 However, the region was extensively modified by mutations with multiple frame shifts and inversions, and probably represented remnants left from the reverse transcriptase gene of a LINE element. It is unlikely that it would be active, due to extensive modifications in the amino acid sequence such that only 18% of the full reverse transcriptase sequence was present and the highly conserved amino acid motif (Y/FXDD) was absent from the sequence 20 (Xiong and Eickbush 1990). The complete sequence was also scanned for the presence of open reading frames (ORFs) that code for proteins using the program GENSCAN (<http://CCR-081.mit.edu/GENSCAN.html>) (Burge and Karlin 1997). Only one gene was found and it corresponded to the *APPA* phytase gene. GENSCAN unexpectedly predicted a different N-terminus for the phytase than would have been expected from the sequence. 25 However, that could have resulted from the lower accuracy of GENSCAN for detecting initiation sites (Burge and Karlin 1998).

3) Generation of Transgenic Mice Expressing a Constitutive Salivary Phytase

In the following description, a pair of founder mice, incorporating the phytase gene and a constitutive promoter, were prepared under contract by the University of Alabama. As 30 will be discussed, these founders were used to produce offspring, which were then analyzed for the presence of the phytase gene by PCR and animals containing the gene were then tested constitutive salivary phytase production.

Two transgenic founder mice (a black male and a white female, 3-1) containing the phytase transgene were received from the NICHD Transgenic Mouse Development Facility at the University of Alabama. The black male was negative for salivary phytase, but the female, 3-1, exhibited a salivary phytase activity of 30 U/ml. Progeny produced by crossing 5 the black male with 4 CD-1 females produced 9 out of 25 females and 13 out of 26 males that were PCR positive. All progeny were negative for salivary phytase. The female founder, 3-1, was out-crossed with a CD-1 male to produce 3 litters for a total of 35 offspring. Of the progeny from these matings one phytase positive G1 male was obtained. When the G1 male was outcrossed with 6 CD-1 females, of the 6 litters 20/34 males were PCR positive and 10 salivary phytase positive and 21/28 females were PCR positive and salivary phytase positive (Table 3). The salivary phytase activity of different offspring from the same first generation (G1) male ranged from 1.3 to 71.2 U/ml. There was no significant difference in the phytase activities between male or female mice.

PCR assays for identification of the transgenic mice were carried out with an initial 15 heating step at 95°C for 3 min, 40 cycles using 95°C for 30 sec, 54°C for 30 sec and 72°C for 1 min) using the following primers: APPA-UP2 and APPA-KPN (Figure 6).

The phytase assays were conducted as described above for the R15-PRP/APPA phytase expressed in cell culture.

20 **4) Production of Transgenic Pigs Containing the Phytase Transgene Lama 2/APPA**

Transgenic pigs were produced using Yorkshire and Yorkshire/Landrace cross gilts as the embryo donors and Yorkshire sows as the recipients. The experimental procedure used was similar to that described by Wall et al. (1985). The detailed procedure is described below. The Lama2/APPA construct with the APPA signal peptide was used as the transgene 25 for microinjection.

Methodology for the generation of transgenic pigs

The following is a description of the preferred method of generating transgenic pigs according to the invention. However, it will be apparent to those skilled in the art that various other methods are also applicable.

30

a) Superovulation of prepuberal gilts and sows.

Selected Yorkshire or Yorkshire/Landrace cross gilts between 70 to 80 kg were superovulated by intramuscular injection of 2000 IU of pregnant mare's serum gonadotropin

(PMSG, Ayerst Veterinary Laboratories), followed by 700 IU human chorionic gonadotropin (HCG, Ayerst Veterinary Laboratories) 60 to 72 hours later, administered in the same manner. The gilts were artificially inseminated three times with a 16 hour interval between inseminations using semen from a high breeding index Yorkshire boar. Twenty-four hours 5 after the last insemination, the gilts were slaughtered and the reproductive tract recovered.

b) Synchronization of estrus in recipients

Estrus was synchronized in experienced recipient sows as described for donor sows. Since synchronization and not superovulation was the goal, hormone levels were reduced to 10 500 IU for PMSG and 500 IU for HCG. PMSG was given the day the sow's litter was weaned, followed in 72 hours by HCG and surgery for embryo transfer was performed 54 hours thereafter.

c) Embryo collection

15 Reproductive tracts were collected at the abattoir, inserted into bags, sealed and the bags immersed in water at 39°C for transport to the laboratory. Recovery of the embryos and microinjection with the transgene was conducted in a laboratory maintained at 32 to 33°C. The oviducts were dissected from the tracts and flushed, using a syringe and a feeding tube, with 15 ml of pre-warmed HBECM-3 medium (Dobrinsky *et al.* 1996). The media was 20 collected in a 100 mm Petri dish and placed in an incubator at 38.5°C with an atmosphere of 5% (vol/vol) of CO₂, 5% (vol/vol) O₂ and the balance N₂. After all tracts were flushed, embryos were individually collected from the flushed media using a polished transfer pipette. Embryos were rinsed twice in 3 ml volumes of pre-incubated BECM-3 and placed in 100 µl 25 of pre-incubated BECM-3 under 3 ml of filter sterilized mineral oil until injected.

25

d) Pronuclear injection

Embryos from one gilt were collected and placed in one ml of pre-warmed HBECM-3 in a 1.5 ml centrifuge tube and centrifuged for 6 min at 14,000 x g (Wall *et al.* 1985). The embryos were then collected and placed in an injection dish with 40 µl of pre-warmed 30 HBECM-3 covered with 2.5 ml of filter sterilized mineral oil. The pronucleus in each embryo was injected (Gordon *et al.* 1980) with three picolitres of Lama2/APPA DNA in solution at a concentration of 5 ng of DNA per µl in 10 mM Tris, pH 7.5, 0.1 mM EDTA. After injection, the embryos were placed in dishes containing 100 µl of pre-incubated

BECM-3 under 3 ml of filter sterilized mineral oil. After all embryos were injected, which took no more than 4 hours since collection of reproductive tracts, the embryos were transferred to 1.8 ml cryotube (Nunc) containing 1 ml of pre-warmed HBECM-3 and transported in an incubator at 38.5°C to the swine surgery.

5

e) Embryo transfer

Recipient sows were anesthetized by intravenous injection of 500 mg Brietol and anesthesia maintained by inhalation of 3% halothane with 4 litres per min of nitrous oxide and 4 litres per min oxygen. The oviducts were exposed through a laparotomy, just off the 10 dorsal midline, and a catheter, containing 20 to 35 injected embryos and 3 to 6 untreated embryos, was passed into the infundibulum and down the oviduct to the isthmus and emptied. The oviduct was returned to the abdominal cavity and the incision closed.

f) Growth of pigs

15 New-born piglets were kept together until weaning. At that time males and females were separated and penned with non-transgenic same sex pigs of a similar age from other litters. The pigs are fed *ad libitum* starter rations until 25 kg wt, grower diet from 25 to 60 kg wt and finisher diet from 60 kg to market weight. Water is available *ad libitum*. Transgenic pigs 167-02, 282-02 and 282-04 were maintained on a low phytate ration until 85, 20 50, and 50 days of age, respectively, and then switched to the grower ration. All other transgenic pigs were given the standard high phosphorus diets.

The diets were provided as pelleted formulations during the weanling, grower and finishing phases are shown in Tables 4 and 5. The vitamin and mineral mixes included in the diets are shown in Tables 6 and 7.

25

PCR analysis

Tail segments from newborn piglets were collected and slices of each placed in 600 µl of 50 mM NaOH and heating at for 95°C for 15 minutes. The suspension was neutralized with 50 µl of 1 M Tris (pH 8.0) and insoluble materials removed by centrifugation for 5 min 30 in a microcentrifuge. A 2 µl sample of each was used for PCR with primers APPA-UP2 and APPA-KPN.

The primers produce a 750 bp fragment if the transgene is present. As a positive control PIG-BGF and PIG-BGR primers were used to detect the porcine β-globin gene from

the same DNA preparation (Heneine and Switzer 1996). The PCR reaction was performed using the same conditions as described for detection of the phytase transgene. As a negative control genomic DNA from a non-transgenic pig was used in the PCR reaction, for a positive control this DNA was spiked with a known amount of transgene (1 gene copy/per genome).

5 When a positive signal was identified by PCR for pig 167-02 (Figure 3) another DNA preparation was made and two more pairs of PCR primers were used to test for gene integrity (Figure 4) APPA-MATURE with APPA-KPN, and APPA-MATURE with APPA-DOWN2 PCR conditions were similar to those described previously.

10 Extraction of DNA from blood for PCR analysis

The method for extraction of DNA from blood was based on a method described by Higuchi (1989) with some modifications. A 100 μ l volume of whole blood was mixed with 200 μ l of lysis buffer (10 mM Tris-HCl, 0.32 M sucrose, 5 mM MgCl₂, 1% (vol/vol) Triton X-100, pH 7.5.), mixed briefly and incubated on ice for 5 min. The sample was then 15 centrifuged at 14,000 x G for 3 min, and the supernate discarded. The sediment was suspended in lysis buffer, mixed, incubated and centrifuged. This procedure was repeated 2 more times, or until no hemoglobin remained. The sediment was dissociated in 100 μ l of 50 mM NaOH, mixed and heated at 100°C for 10 min. The contents were cooled, 10 μ l of 1 M Tris-HCl (pH 8.5) added and mixed briefly. The sample was then centrifuged at 14,000 x g 20 for 2 min and 2 μ l of the supernate used for analysis by PCR.

The PCR reaction mixture with a total volume of 40 μ l consisted of, 23.8 μ l of distilled water, 4 μ l of 10 X Gibco BRL PCR buffer, 1.2 μ l of 50 mM MgCl₂, 0.8 μ l of 10 mM dNTPs, 40 pmol of each of the forward and reverse primers in 8 μ l, 2 μ l of template DNA and 0.2 μ l of *Taq* DNA polymerase (Gibco BRL, 5 U/ μ l). The amplification procedure 25 was performed with an initial heating step at 95°C for 3 min followed by 40 cycles of 95°C for 30 sec, 54°C for 30 sec and 72°C for 60 sec.

The transgenic pigs were detected with primers for the APPA gene (APPA-KPN with APPA-UP2), and as a control PIG-BGF with PIG-BGR primers were used for detection of the porcine β -globin gene.

30

Saliva collection from pigs for phytase assays and weighing of pigs

Weanling pigs were sampled for salivary phytase by wiping under the tongue with a cotton tipped applicator, breaking the stick off and centrifuging the applicator tip in a 0.4 ml

microcentrifuge tube, with a hole in the bottom, contained within a 1.5 ml microcentrifuge tube. Grower and finishing pigs were sampled using 1.5 inch long #2 dental cotton absorbent rolls (Ash Temple Sundries Ltd, Don Mills, ON) attached to dental floss. These were centrifuged in 1.5 ml microcentrifuge tubes with holes in the bottom while contained in larger tubes. The saliva was collected from the larger tube and stored at -20°C until analyzed.

5 Saliva was collected and pigs were weighed at weekly intervals.

Analysis for phytase activity.

10 Saliva samples were either assayed directly or after dilution in 0.1 M acetate buffer pH 4.5. Phytase was assayed in 200 µl of 0.1 M sodium acetate buffer (pH 4.5) using sodium phytate (4 mM) as a substrate at 37°C. After 10 min of incubation the reaction was stopped by addition of 133 µl ammonium molybdate/ammonium vanadate/nitric acid mixture and the concentration of liberated inorganic phosphate determined at 405 nm (Engelen, van der Heest, Randsdorp, and Smit 1994). This and all other assays were performed in triplicate.

15 One unit (U) of enzyme activity was the amount of the enzyme releasing 1 µmol of inorganic phosphate per minute.

Assays for salivary phytase and for phytase in blood samples were conducted as previously described for saliva samples. A reagent blank with blood added at the same concentration as the samples assayed was subtracted from the sample readings.

20

Collection of fecal materials and analysis for total phosphorus

Fresh feces were collected from each pig during the grower and finisher phases. Samples were placed in aluminum trays closed with a wax paper top and immediately frozen, and kept frozen until they were lyophilized for analysis. After lyophilization the samples 25 were transferred to room conditions overnight to reach equilibrium in moisture content. The samples were separately ground with a mortar and pestle until homogenous and sealed in plastic containers until analyzed further. Dry matter content of samples was analyzed according to AOAC (Association of Official Analytical Chemists (AOAC) 1984) by heating 1 gram samples at 110°C for 4 hours and cooling in a desiccator prior to weighing. To 30 analyze total phosphorus content, samples were heated at 550°C in a muffle furnace and 10 ml of 10 M HCl added and heated to boiling. The contents from each sample was quantitatively diluted to 250 ml with water and inorganic phosphorus content was measured by the method of Heinoen and Lahti (1981).

Purification of the *E. coli* produced phytase and pig salivary phytase

The APPA phytase was over expressed in *E. coli* strain BL21(DE3) and the EDTA lysozyme extract fraction purified on DEAE-Sepharose and Sephadex-G75 as described by 5 Jia et al. (1998). The pig phytase was purified by chromatography on DEAE-Sepharose and the band of enzyme eluted with a sodium chloride gradient was further purified by Chromatofocusing using a pH gradient from pH 4.0 to 7.0.

SDS-PAGE analysis and Silver Staining

10 Sodium dodecylsulfate polyacrylamide gel electrophoresis was performed using a 10% gel as described by Laemmli (1970), except that protein in the sample buffer was heated at 70°C for 10 minutes. Samples were stained with silver as described by Nesterenko et al. (1994).

15 Preparation of a monoclonal antibody specific for the APPA encoded *E. coli* phytase

Monoclonal antibodies specific to the *E. coli* APPA encoded phytase were prepared according to the procedures of Galfrè and Milstein (1981). Briefly, two female Balb/c mice were immunized 7 times over a period of 59 days with a purified APPA enzyme preparation. Mouse spleens were harvested, and the cells therein fused with an NS-1 myeloma cell line 20 (Kohler and Milstein, 1976). Fused cells were selected for their ability to grow in media containing hypoxanthine, aminopterin, and thymidine (HAT). Western blotting and Enzyme-Linked Immunosorbent Assays (ELISA) were used identify those clones capable of secreting an antibody into the culture medium that recognized epitopes on both the *E. coli* and pig derived APPA enzyme. Clones secreting a desirable antibody were subcloned twice to 25 ensure a pure culture of antibody secreting hybridomas.

Production of Polyclonal Antibodies Against the Purified *E. coli* derived APPA Phytase

Antibodies were prepared in two New Zealand White Rabbits by two intramuscular injections at different sites in the thigh of 50 µg of purified *Escherichia coli* derived APPA 30 phytase in 0.5 ml of a 1:1 mixture of phosphate-buffered saline (PBS) and Freund's Complete Adjuvant. This was followed by repeat injections of 20 µg each of phytase in a 1:1 mixture of PBS and Freund's Incomplete Adjuvant on days 4, 19, 25, and 39. Blood was collected via heart puncture on day 42. The serum was separated from the cell fraction and used as the

source of antibodies. The basic procedures for antibody production are described in Harlow and Lane (1988).

Western blotting

5 Western blotting was performed as described by Towbin *et al.* (Towbin *et al.* 1979).
Deglycosylation of pig phytase was done according to protocols, Roche Molecular
Biochemicals, with following modifications. Protein in 50 mM Tris (pH 8.0), 1 mM EDTA,
1% SDS, 1% 2-mercaptoethanol was denatured by heating at 95° C for 3 min. Then protein
was precipitated with chloroform-methanol method (Wessel and Flugge 1984) and
10 resuspended at 100 µg/mL in 20 mM Sodium Phosphate (pH 7.2) with 1% Triton X-100.
Complete deglycosylation of 5 µg in 50 µL phytase was carried out overnight at 37°C using
1 unit (U) N-glycosidase F, 1.2 mU O- glycosidase and 1 mU neuraminidase (Boehringer
Mannheim GmbH). After incubation 0.5 µg of protein was run on the SDS gel.

15 Staining of glycoproteins

This staining was done using DIG Glycan Detection Kit (Boehringer Mannheim)
according to manufacture instructions (O'Shannessy *et al.* 1987).

Statistics on the generation of transgenic pigs

20 The statistics on embryos recovered, microinjected and transferred into donor sows is
shown in Table 8. A total of 4147 embryos injected with the transgene and 675 untreated
embryos were introduced into 140 recipient sows with an average of 30 injected embryos and
5 uninjected embryos. All offspring were tested for the presence of the transgene in tissue
biopsy, in blood by PCR analysis, and by an assay for phytase activity in the saliva.

25 Table 9 lists the transgenic pigs that were produced, their birth dates, sex and salivary
phytase levels. There were 31 pigs transgenic for the phytase gene out of 203 live piglets
born from embryos microinjected. These were detected by the presence of the gene in blood
samples using the standard primer set, APPA-UP2 and APPA -KPN, but only 14 were
detected by analysis of tail DNA preparations using the standard primer set. When the
negative samples were reanalyzed using the primer set LAMA-UP1 and APPA-down4
30 (Figure 8) a further 8 tail DNA samples were found to be positive. Purification of the tail
biopsy DNA probably would have led to all being PCR positive for the phytase transgene.

Characteristics of the phytase transgene in transgenic pig 167-02

The application of PCR to detection of transgenic pigs is exemplified by analysis of litter 167 in which one of 7 piglets tested, including one that was stillborn and one that was crushed by the sow after birth, one live piglet designated 167-02 was identified as positive for 5 the APPA gene by generation of a PCR product (Lane 2) of approximately 750 bps from the tail chromosomal DNA (Figure 7). No rearrangements of the APPA gene were detected as documented by the positive PCR results using primers directed to the 3' region (lane 2) the whole gene (lane 3) and the 5' region (lane 4) of the APPA gene (Figure 8).

10 Salivary phytase and weight gain during growth of transgenic and non-transgenic penmates.

Data on salivary phytase activity and weight gain are shown for five transgenic pigs and for weight gains of their non-transgenic penmates in Figures 9, 10, 11, 12 and 13. The phytase activity in the saliva varied substantially from one sampling time to the next. This variability was attributed to a combination of environmental factors including whether the 15 animal had just consumed food or water, and regulation of parotid and saliva secretion in relation to food and water consumption. The weight gains during growth of the five transgenic pigs was within the range of the weight gains of the normal non-transgenic pigs.

With the exception of 167-02 the growth rate of the transgenic pigs was similar to that of the non-transgenic litter mates.

20 Phosphorus content in the fecal materials from transgenic and non-transgenic pigs.

The phosphorus content of fresh fecal samples from three of the transgenic founder pigs, 167-02, 282-02, 282-04, 405-02 and 421-06 receiving weaning, grower or finisher ration is shown in Table 9. The phosphorus content of the feces of the transgenic pigs ranged from 1.59 to 2.26% while that of the non-transgenic penmates ranged from 1.61 to 2.76 %. 25 The reduction in fecal phosphorus ranged from a maximum of 26% to a minimum of 8%. In most cases the differences were at the 99% level of significance. The ages of the pigs at the time of fecal sampling and the corresponding phytase activities are shown in Figures 9, 10, 11, 12 & 13. The rations fed contained a supplement of readily available phosphorus suitable for maximizing growth of non-transgenic pigs. Since the reduction in fecal phosphorus is 30 measured in transgenic pigs receiving a diet high in mineral phosphorus it is very likely that the fecal phosphorus would be substantially lower if the diet lacked mineral phosphorus. Under these conditions the phosphorus released from phytate would provide a substantial

proportion of the dietary phosphorus and little would reach the large intestine and be excreted in the feces.

Transmission of the phytase transgene (to be completed)

5 When semen from the transgenic boar 167-02 was used to inseminate four Yorkshire gilts all four sows had litters in which 4 out of 8, 2 out of 9, 7 out of 8 and 2 out of 5 of the piglets were transgenic for the phytase gene (Table 11). The PCR data for litter 154 that documents the presence of the transgene is shown in Figure 14. All pigs containing the gene exhibited phytase activity in the saliva, and it ranged from 341 to 10,077 units per ml. Half
10 of the transgenic piglets had salivary phytase activities of greater than 2000 units per ml. The specific activity of the phytase in the saliva ranged from 39 U/mg protein to a high of 706 units/mg protein.

15 This data documents that the gene was transferred and that the level of phytase expression observed in the founder was preserved in the first generation of pigs. Both male and female pigs at 11 days of age exhibited high phytase activity.

Characteristics of the phytase enzyme synthesized in the salivary glands of the pig

20 The phytase enzyme was purified to homogeneity from *E. coli* and from saliva collected from transgenic pig 167-02. Silver stains of the purified enzymes after SDS-PAGE are shown in Figure. 15. The *E. coli* derived enzyme has a molecular mass of approximately 45 kDa while that produced by the pig was about 55 kDa. The enzymes were also electrophoresed as before, transferred to nitrocellulose and stained for glycoproteins. The second part of Figure 15 shows that the pig APPA protein is glycosylated. Figure 15B shows that treatment of the pig phytase with deglycosylation enzymes changes the size of the
25 phytase from 60 kDa to 45 kDa, an observation that confirms the glycosylated nature of the recombinant phytase produced in the saliva of the pig.

The data in Figure 16 shows that the pig phytase is homologous with the *E. Coli* enzyme despite their difference in size.

30 The purified pig phytase had K_m and V_{max} values of 0.33 mM and 624 units per mg of protein, respectively. Golovan et al. (2000) previously reported the K_m and V_{max} for the *E. coli* enzyme to be 0.63 mM and 2325 units per mg of protein. Thus the salivary phytase exhibits approximately 25% of the activity of the *E. coli* enzyme. This reduction in activity may be due to glycosylation that either modifies the catalytic site of the enzyme or otherwise leads to the formation of an enzyme with lower catalytic activity.

The latter finding of the production of a glycosylated protein suggests a method of producing such proteins using transgenic animals. Currently, although recombinant methods are available for producing proteins in host cells, it is often found that the mature peptide lacks the glycosylation normally associated with proteins produced by higher life forms.

5 Insulin is an example of such protein. The findings of this study suggest that one means of producing the desired glycoproteins would be to generate transgenic animals such as the pig, that have been transformed, by known methods or the method described above, with a gene encoding the desired protein. When expressed by such animal, the subject protein would be produced and would undergo post-translational processing in the cell including the step of
10 glycosylation. Thus, the invention contemplates a general method of producing such glycosylated proteins. Further, the invention contemplates a method of producing glycosylated proteins through the expression in and isolation from the saliva of an animal that has been transformed with a gene encoding such protein, and wherein such gene is operably linked to a saliva protein promoter or enhancer.

15 Various methods are known in the art for the collection of glycoproteins from the parotid gland of the pig for various applications. For example, surgical techniques have been published by Denny et al. (1972) for the collection of secretions from the parotid gland and submandibular salivary ducts.

20 Test kit for detection of the APPA phytase protein in pigs

The monoclonal antibodies produced against the APPA phytase expressed in *E. coli* reacted with the APPA phytases produced in the saliva of transgenic mice and pigs (Figure 17). Immunological detection of phytase in saliva provides definitive proof that the phytase secreted in transgenic pig saliva is a product of the *APPA* gene expressed in the pig salivary
25 gland. This serves as a reliable method to document phytase production in transgenic pigs.

A further test would also be obtainable using the polyclonal antibodies discussed above.

30 The DNA sequence encoding phytase may be obtained from a variety of sources such as microbial, plant or animal sources. Preferably, the DNA sequence is obtained from a microbial source such as bacteria. Most preferred DNA sequences are obtained from *Escherichia coli*.

The cloning of a gene or a cDNA encoding a phytase protein may be achieved using various methods. One method is by purification of the phytase protein, subsequent

determination of the N-terminal and several internal amino acid sequences and screening of a genomic or cDNA library of the organism producing the phytase using oligonucleotide probes based on the amino acid sequences. If at least a partial sequence of the gene is known, this information may be used to clone the corresponding cDNA using, for instance, the 5 polymerase chain reaction (PCR) (PCR Technology: Principles and Applications for DNA Amplification, (1989) H. A. Ehrlich, ed., Stockton Press, New York; the contents of which are incorporated herein by reference). It will be evident to those skilled in the art that the cloned phytase gene described above may be used in heterologous hybridization experiments, directed to the isolation of phytase encoding genes from other microorganisms.

10 The DNAs encoding phytase or individual fragments or modified proteins thereof can be fused, in proper reading frame, with appropriate regulatory signals as described in detail below, to produce a genetic construct that is then amplified, for example, by preparation in a bacterial (e.g., *E. coli*) plasmid vector according to conventional methods. Such methods are described in, for example, Sambrook et al., Molecular Cloning: A Laboratory Manual (Cold 15 Spring Harbor Press 1989), the contents of which are incorporated herein by reference. The amplified construct is thereafter excised from the vector and purified for use in producing transgenic animals.

20 The desired protein may also be produced as a fusion protein containing another protein. For example, the desired recombinant protein of this invention may be produced as 25 part of a larger recombinant protein in order to stabilize the desired protein. Useful modifications within this context include, but are not limited to, those that alter post-translational modifications, size or active site, or that fuse the protein or portions thereof to another protein. Such modifications can be introduced into the protein by techniques well known in this art, such as by synthesizing modified genes by ligation of overlapping oligonucleotides or introducing mutations into the cloned genes by, for example, oligonucleotide-mediated mutagenesis.

30 The cloned phytase gene may be used as starting materials for the construction of improved phytases. Improved phytases are phytases, altered by mutagenesis techniques (e.g. site-directed mutagenesis, or directed evolution), which have properties that differ from those of wild-type phytases (Kuchner and Arnold 1997). For example, the temperature or pH optimum, specific activity, temperature or protease resistance may be altered so as to be better suited for a particular application.

A choice of expression in cellular compartments (such as cytosol, endoplasmic reticulum) or extracellular expression can be used in the present invention, depending on the

biophysical and biochemical properties of the phytase. Such properties include, but are not limited to pH sensitivity, sensitivity to proteases, and sensitivity to the ionic strength of the preferred compartment. The DNA sequence encoding the enzyme of interest should be modified in such a way that the enzyme can exert its action at the desired location in the cell.

5 To achieve extracellular expression of the phytase, the expression construct of the present invention utilizes a bacterial signal sequence. Although signal sequences that are homologous (native) to the animal host species are preferred, heterologous signal sequences, i.e. those originating from other animal species or of microbial origin, may be used as well. Such signal sequences are known to those skilled in the art.

10 All parts of the relevant DNA constructs (promoters, regulatory, secretory, stabilizing, targeting, or termination sequences) of the present invention may be modified, if desired, to affect their control characteristics using methods known to those skilled in the art. The *cis*-acting regulatory regions useful in the invention include the promoter that drives expression of the phytase gene. Highly preferred are promoters that are specifically active in salivary gland cells. Among such promoters, highly preferred are mouse parotid secretory protein (PSP) promoter, rat proline-rich protein (PRP) promoter, human salivary amylase promoter, mouse mammary tumor virus promoter (Samuelson 1996). Among the useful sequences that regulate transcription, in addition to the promoters discussed above, are enhancers, splice signals, transcription termination signals, and polyadenylation sites. Particularly useful in this regard are those that increase the efficiency of the transcription of the genes for phytase in the salivary gland or other cells of the transgenic animals listed above. Preferred are transcription regulatory sequences for proteins highly expressed in the salivary gland cells. Introns could be introduced to increase levels of expression. Such introns include the synthetic intron SIS, SV40 small t antigen intron and others (Whitelaw *et al.* 1991; Petitclerc *et al.* 1995).

25 Preferably, the expression system or construct of this invention also includes a 3' untranslated region downstream of the DNA sequence encoding the desired recombinant protein, or the salivary protein gene used for regulation. This region apparently stabilizes the RNA transcript of the expression system and thus increases the yield of the desired protein.

30 Among the 3' untranslated regions useful in this regard are sequences that provide a polyA signal. Such sequences may be derived, e.g., from the SV 40 small t antigen late polyadenylation signal, synthetic polyadenylation signal or other 3' untranslated sequences well known in this art (Carswell and Alwine 1989; Levitt *et al.* 1989). Preferably, the 3' untranslated region is derived from a salivary-specific protein. The stabilizing effect of this

region's polyA transcript is important in stabilizing the mRNA of the expression sequence. Further, the addition of locus control regions (LCRs), matrix attachment regions (MAR) and scaffold attachment regions (SARs) would allow position-independent, copy number dependent expression of the transgene with either homologous or heterologous promoters

5 (Taboit-Dameron *et al.* 1999; Geyer 1997). Co-integration of an actively expressed gene with the transgene was also shown to increase expression levels of a poorly expressed transgene (Clark *et al.* 1993). Also important in increasing the efficiency of expression of phytase is a strong translation initiation site (Kozak 1987). Likewise, sequences that regulate the post-translational modification of phytase may be useful in the invention.

10 The term "animal" as used herein denotes all animals except humans. It also includes an individual animal in all stages of development, including embryonic and fetal stages.

A "transgenic" animal is any animal containing cells that bear genetic information received, directly or indirectly, by deliberate genetic manipulation at the subcellular level, such as by microinjection or infection with a recombinant virus. "Transgenic" in the present

15 context does not encompass classical crossbreeding or in vitro fertilization, but rather denotes animals in which one or more cells receive a recombinant DNA molecule. Although it is highly preferred that this molecule be integrated within the animal's chromosomes, the invention also encompasses the use of extrachromosomally replicating DNA sequences, such as might be engineered into yeast artificial chromosomes. The information to be introduced

20 into the animal may be foreign to the species of the animal to which the recipient belongs (i.e., "heterologous"), or the information may be foreign only to the particular individual recipient, or genetic information already possessed by the recipient. In the last case, the introduced gene may be expressed in a manner different than the native gene.

As indicated above, the transgenic animals of this invention are other than human.

25 Farm animals (pigs, goats, sheep, cows, horses, rabbits and the like), rodents (such as mice and rats), domestic pets (e.g. cats and dogs), fish and poultry (e.g. chickens) are included in the scope of this invention. It is highly preferred that a transgenic animal of the present invention be produced by introducing into single cell embryos appropriate polynucleotides that encode phytase, or fragments or modified products thereof, in a manner such that these

30 polynucleotides are stably integrated into the DNA of germ line cells of the mature animal, and are inherited in normal mendelian fashion. Advances in technologies for embryo micromanipulation now permit introduction of heterologous DNA into fertilized mammalian ova. For instance, totipotent or pluripotent stem cells can be transformed by microinjection, calcium phosphate mediated precipitation, liposome fusion, retroviral infection or other

means, the transformed cells are then introduced into the embryo, and the embryo then develops into a transgenic animal. In one preferred method, developing embryos are infected with a retrovirus containing the desired DNA, and transgenic animals produced from the infected embryo. In a most preferred method, however, the appropriate DNAs are co-injected 5 into the pronucleus or cytoplasm of embryos, preferably at the single cell stage, and the embryos allowed to develop into mature transgenic animals. Such techniques are well known (see reviews of standard laboratory procedures for microinjection of heterologous DNAs into mammalian fertilized ova, including Hogan et al., Manipulating The Mouse Embryo, (Cold Spring Harbor Press 1986); Krimpenfort et al., Bio/Technology 9:844 (1991); Palmiter et al., 10 Cell, 41: 343 (1985); Kraemer et al., Genetic Manipulation Of The Early Mammalian Embryo, (Cold Spring Harbor Laboratory Press 1985); Hammer et al., Nature, 315: 680 (1985); Wagner et al., U.S. Pat. No. 5,175,385; Krimpenfort et al., U.S. Pat. No. 5,175,384, the respective contents of which are incorporated herein by reference).

For a person skilled in art, it will also be clear that the present invention provides for 15 other proteins to be expressed in the salivary gland of the pig. Such proteins may be secreted into saliva to improve digestion and decrease pollution potential (for example, endoglucanases), or specifically targeted for secretion into blood and have effects on the growth and health of the animal (such as growth hormone).

Phytase activity may be measured via a number of assays, the choice of which is not 20 critical to the present invention. For example, the phytase enzyme activity of the transgenic animal tissue may be tested with an ELISA-assay, Western blotting or direct enzyme assays using calorimetric techniques or gel assay system.

The examples included herein are provided so as to give those of ordinary skill in the art a complete disclosure and description of how to make and use the invention and are not 25 intended to limit the scope of what the inventors regard as their invention. Efforts have been made to ensure accuracy with respect to numbers used (e.g., amounts, temperature, pH, etc.) but some experimental errors and deviation should be accounted for. Unless indicated otherwise, temperature is in degrees Centigrade and pressure is at or near atmospheric.

30

Although the invention has been described with reference to certain specific embodiments, various modifications thereof will be apparent to those skilled in the art without departing from the spirit and scope of the invention as outlined in the claims appended hereto.

Table 1. Secretion of phytase in the saliva of transgenic mice containing the R15-PRP/APPA transgene and non-transgenic mice induced with isoproterenol and pilocarpine.

Founder	Mice	PCR	Gender	Generation	Transgene	Phytase activity micromoles/min/ml
A0m	4bfr (+)	positive	F	1	APPA+intron	39.73
A0m	2brm(+)	positive	M	1	APPA+intron	24.29
A0m	2brm(+)	positive	M	2	APPA+intron	14.42
A0m	5brf(+)	positive	F	2	APPA+intron	7.36
A0m	1brm(-)	negative	M	1	APPA+intron	0.00
A1f	9brf(+)	positive	F	1	APPA+intron	0.08
A1f	11w f(+)	positive	F	1	APPA+intron	0.07
A1f	5brm(+)	positive	M	1	APPA+intron	0.03
A1f	10wf(-)	negative	F	1	APPA+intron	0.02
A20f	1brm(+)	positive	M	1	APPA+intron	0.53
A20f	5brf(+)	positive	F	1	APPA+intron	0.12
A20f	4brf (-)	negative	F	1	APPA+intron	0.03
A2m	13wf(+)	positive	F	1	APPA+intron	87.70
B0m	5brf (+)	positive	F	1	APPA+intron	0.95
B0m	3brm(+)	positive	M	1	APPA+intron	0.73
B0m	6wf (-)	negative	F	1	APPA+intron	0.00
B0f	3wf (+)	positive	F	2	APPA	252.43
B0m-intr	9wf(+)	positive	F	1	APPA	546.74
W0m	8wf(+)	positive	F	1	APPA	60.42
W30m	1wm(+)	positive	M	2	APPA	41.91
W30m	11w f(+)	positive	F	1	APPA	43.44
W30m	4wm(-)	negative	M	1	APPA	0.02
W30m	10wf (-)	negative	F	1	APPA	0.02

Table 2. Repeat sequences found in the Lama2-APPA construct.

Start	End	DNA strand	Repeat	Class/family	Substitutions % of consensus	Deletions % of consensus	Insertions % of consensus
765	927	+	L1M1	LINE/L1	25	4.2	6.7
928	965	+	(CA)n	Simple repeat	0	0	0
966	1020	+	L1M1	LINE/L1	25	4.2	6.7
1021	1156	+	B1_MM	SINE/Alu	15.4	0	0
1159	1231	+	CAAC)n	Simple repeat	1.4	0	0
1232	1385	+	L1M1	LINE/L1	25	4.2	6.7
1652	2308	C	L1	LINE/L1	28.5	11.9	1.7
2334	2406	C	MIR	SINE/MIR	27.4	4.1	0
2415	3266	+	RMER13A	LTR	17.7	4	6.1
6016	6127	C	L1MA9	LINE/L1	25.5	2	1
6831	7007	+	CT-rich	Low complexity	30.5	1.7	3.4
7299	7510	C	B3	SINE/B2	27.8	7.5	1.4
7718	7746	+	(TCTCTG)n	Simple repeat	6.9	0	0
8499	8581	C	MIR	SINE/MIR	24.1	12.1	3.6
9010	9603	+	Lx4	LINE/L1	21.7	6.4	0.2
10465	10519	+	(TG)n	Simple repeat	5.5	1.8	0
11235	11287	C	MERSA	DNA/MER1 type	28.3	0	1.9
12372	12537	C	L1MA4A	LINE/L1	28.3	5.4	0
14240	14388	+	B1_MM	SINE/Alu	4	0	1.3
14869	14945	C	MIR	SINE/MIR	36.4	1.3	0
16391	16540	C	ORR1D	LTR/MaLR	29.3	0	6
16774	17214	+	RMER4	LTR	21.3	10	11.8
17229	17718	C	L1_MM	LINE/L1	15.3	0	0.8

Table 3. Salivary phytase activities of G2 mice from the founder female 3-1 generated using the construct Lama2-APPA. The mice were between 21 and 30 days of age.

male mouse #	Phytase (U/ml)	female mouse #	Phytase (U/ml)
5	28.3	1	9.0
6	2.5	2	29.9
8	6.6	4	8.0
9	44.7	5	43.0
10	12.7	6	26.9
12	28.3	8	1.9
15	28.1	9	66.3
18	71.2	10	19.9
19	19.5	11	61.3
20	15.7	12	36.4
21	20.9	13	18.0
22	4.1	17	38.9
24	13.0	18	18.5
26	53.4	19	27.0
28	20.4	23	6.5
29	34.1	24	16.1
30	11.1	25	9.4
32	3.1	26	14.8
33	51.7	27	1.3
34	19.0	28	8.2

Table 4. Composition and nutrient levels of Phase II starter diet and low phytate starter diets fed to weanling pigs between 5-10 kg.

Ingredients	Diet/Nutrient Levels ¹	
	Phase II Starter Diet	Low Phytate Starter Diet
Corn	33.15	25.44
Barley	8.00	8.00
Wheat	20.00	40.00
Soybean meal	21.00	8.00
Fish meal	5.00	5.00
Meat and bone meal	-	1.00
Whey	8.00	8.00
Fat	2.00	2.00
Lysine-HCl	0.10	0.28
Dicalcium phosphate	1.10	-
CaCO ₃	0.90	1.10
Iodized salt	0.30	0.30
Vitamin premix ¹	0.250	0.55
Mineral premix ¹	0.10	0.10
Lincommix 44	0.10	0.10
Total (kg)	100.00	100.00
Calculated nutritive values		
DE (kcal/g)	3.44	3.36
CP (%)	19.46	18.62
Ca (%)	1.00	0.94
Total P (%)	0.74	0.66
Ca/P	1.35:1	1.42:1
Total AA contents (%)		
Arginine	1.16	1.17
Histidine	0.50	0.48
Isoleucine	0.81	0.77
Leucine	1.58	1.54
Lysine	1.17	1.06
Methionine	0.34	0.29
Cysteine	0.34	0.34
Methionine+Cysteine	0.68	0.63
Phenylalanine	0.90	0.90
Tyrosine	0.65	0.65
Threonine	0.75	0.68
Tryptophan	0.23	0.23
Valine	0.91	0.86

¹Minerals and vitamins meet or exceed levels recommended by NRC (1998).

Table 5. Composition and nutrient levels of grower and finisher diets.

Ingredients	Diet/Nutrient Levels	
	Grower Diet For pigs 20 to 50 kg	Finishing Diet For pigs 50 to 120 kg
Corn	51.78	40.00
Barley	8.10	23.03
Wheat	20.00	23.00
Soybean meal	16.00	13.00
Fat	1.00	1.00
Lysine-HCl	0.12	0.12
Dicalcium phosphate	1.20	1.00
CaCO ₃	1.15	1.15
Iodized salt	0.50	0.50
Vitamin premix ¹	0.15	0.15
Mineral premix ¹	0.10	0.10
Total (kg)	100.00	100.05
Calculated nutritive values		
DE (kcal/g)	3.39	3.33
CP (%)	14.76	14.17
Ca (%)	0.79	0.74
Total P (%)	0.57	0.53
Ca/P	1.39:1	1.39:1
Total AA contents (%)		
Arginine	0.86	0.80
Histidine	0.38	0.36
Isoleucine	0.58	0.55
Leucine	1.28	1.18
Lysine	0.78	0.73
Methionine	0.24	0.23
Cysteine	0.29	0.29
Methionine+Cysteine	0.53	0.52
Phenylalanine	0.70	0.68
Tyrosine	0.50	0.46
Threonine	0.52	0.49
Tryptophan	0.17	0.16
Valine	0.68	0.65

¹Minerals and vitamins meet or exceed levels recommended by NRC (1998).

Table 6. Vitamin premix composition¹

Nutrient	Amount per 5 kg of premix
Wheat midds	3.867 kg
Vitamin A	10 million IU
Vitamin D	1 million IU
Vitamin E	40 thousand IU
Menadione	2.5 g
Pantothenic acid	15 g
Riboflavin	5 g
Folic acid	2 g
Niacin	25 g
Thiamin	1.5 g
Pyridoxine	1.5 g
Vitamin B ₁₂	25 mg
Biotin	200 mg
Choline	500 g

¹From Hoffman-LaRoche Limited, P.O. Box 877, Cambridge, ON. N1R5X9

Table 7. Composition of the mineral premix^{1,2}

Mineral component	Amount (%)
Limestone	43.3
Copper sulfate (25%)	6.0
Ferrous sulfate (30%)	33.4
Zinc oxide (72%)	13.9
Manganous oxide (56%)	3.4

¹Mineral premix prepared at Arkell

²Dicalcium phosphate contained 18.5% calcium and 20.5% of phosphate and normally is added at a level of 1.2% to the pig grower diet, 1.0% to the finisher diet and 1.5% to the nursing sow diet.

Table 8. Statistics on embryo recovery and the introduction of embryos containing the transgene into recipient sows.

Treatment	Number
Gilts used for embryo recovery:	
Yorkshire	279
Yorkshire x Landrace cross	168
Duroc	12
Total	459
Recipient sows ¹	74
Embryos transferred to recipients:	
Embryos microinjected with the transgene	4147
Uninjected carrier embryos	675
Total	4543
Total number of embryo transfers	140

¹Sows were used for up to three farrowings of potentially transgenic pigs. Sows were inseminated with Yorkshire semen from a high breeding value boars.

Table 9. Transgenic pigs containing a salivary phytase gene generated by microinjections of single cell zygotes using the Lama2-APPA transgene

ID # of pig ¹	Birth Date	Presence of Transgene ² Tail/Blood	Sex	Salivary phytase (U/ml) ³	Zygote source ⁴
167-02	Apr 14/99	+/+	Boar	6,000	Yorkshire
282-02	Jun 14/99	+/+	Boar	618	Yorkshire
282-04	Jun 14/99	+/+	Boar	1,349	Yorkshire
405-02	Aug 14/99	+/+	Gilt	339	York/Landrace
421-02	Aug 24/99	-/+	Gilt	0.8	York/Landrace
421-04	Aug 24/99	-/+	Gilt	2.2	York/Landrace
421-06	Aug 24/99	+/+	Boar	97	York/Landrace
448-01	Sep 03/99	+/+	Gilt	0	York/Landrace
491-01	Sep 25/99	+/+	Gilt	2.3	York/Landrace
491-02	Sep 25/99	+/+	Gilt	0	York/Landrace
491-03	Sep 25/99	+/+	Gilt	0.3	York/Landrace
491-05	Sep 25/99	+/+	Boar	0	York/Landrace
496-05	Sep 26/99	+/+	Boar	0	York/Landrace
500-03	Sep 28/99	+/+	Boar	136	York/Landrace
510-01	Sep 28/99	+/+	Boar	0.2	York/
559-05	Nov 01/99	+*/+	Boar	>418	York/Landrace
560-04	Nov 02/99	+*/+	Boar	5	Yorkshire
594-03	Nov 18/99	+/+	Gilt	2.3	Yorkshire
613-02	Nov 27/99	-/+	Gilt	0.5	York/Landrace
613-03	Nov 27/99	-/+	Gilt	0.3	York/Landrace
647-01	Dec 13/99	-/+	Gilt	0.5	York/Landrace
647-03	Dec 13/99	+*/+	Gilt	16.3	York/Landrace
647-04	Dec 13/99	-*/+	Gilt	0.5	York/Landrace
647-08	Dec 13/99	-*/+	Boar	0.4	York/Landrace
647-09	Dec 13/99	+*/+	Boar	1.92	York/Landrace
668-01	Dec 17/99	+*/+	Gilt	489	Yorkshire
671-02	Dec 19/99	+*/+	Boar	6.9	York/Landrace
671-04	Dec 19/99	+*/+	Boar	325	York/Landrace
675-03	Dec 21/99	-*/+	Gilt	2.1	York/Landrace
675-04	Dec 21/99	+*/+	Boar	42.6	York/Landrace
675-06	Dec 21/99	-*/+	Boar	5.0	York/Landrace

¹The number preceding the dash represents the litter number and the number following the dash is the pig number within the litter.

²All PCR assays were conducted with the primer APPA-up2-APPA-Kpn. Assays indicated with a star gave a negative result with the primer pair. However these samples gave a positive result for the primer set APPA-d4-Lama-up1. Samples 613-02 and 613-03 were negative with the latter primer set.

³Saliva was sampled and assayed for phytase 2 to 4 days after birth of the piglets.

⁴Zygotes used for microinjection were collected from superovulated Yorkshire or Yorkshire-Landrace cross gilts.

Table 10. Phosphorus content of feces collected from pigs producing a salivary phytase and non-transgenic pen-mates¹. The data was subjected to a T-test analysis and the data recorded below.

	Mean Fecal Phosphorus (%)	SE	Relative reduction in fecal phosphorus (%)	t	t (1%)
1. 167-02 Grower Diet (122 days):	1.59		24.47		-
Non-transgenic (n=4)	2.11	0.0604669		8.517	4.6
2. 167-02 Finisher Diet (154 days):	1.97		16.97		
Non-transgenic (n=4)	2.37	0.0240767		16.717	4.6
3. 282-02 Grower Diet (93 days):	1.85		12.90		
Non-transgenic (n=5)	2.124	0.022231964		12.324	4.03
4. 282-02 Finisher Diet (145 days):	1.76		16.03		
Non-transgenic (n=5)	2.096	0.099153384		3.389	4.03 ²
5. 282-04 Grower Diet (93 days):	1.95		8.19		
Non-transgenic (n=5)	2.124	0.022231964		7.827	4.03
6. 282-04 Finisher Diet (145 days):	1.56		25.57		
Non-transgenic (n=5)	2.096	0.099153384		5.406	4.03
7. 421-06 Starter II Diet (40 days):	1.17		27.47		
Non-transgenic (n=5)	1.612	0.086155741		5.140	4.03
8. 421-06 Start III Diet (48 days):	1.57		18.01		
Non-transgenic (n=5)	1.915	0.102884789		3.351	4.03
9. 421-06 Grower Diet (81 days):	2.00		13.28		
Non-transgenic (n=5)	2.310	0.151658823		2.022	4.03
10. 421-06 Finisher Diet (136 days):	1.71		21.20		
Non-transgenic (n=5)	2.173	0.053023237		8.687	4.03
11. 405-02 Starter II Diet (40 days):	1.81		26.97		
Non-transgenic (n=5)	2.482	0.173625623		3.856	4.03
12. 405-02 Starter III Diet (48 days):	1.54		36.58		
Non transgenic (n=4)	2.430	0.104642248		8.496	4.6
13. 405-02 Grower Diet (80 days):	2.26		18.19		
Non-transgenic (n=4)	2.763	0.124724697		4.029	4.6
14. 405-02 Finisher Diet (136 days):	2.26		13.24		
Non-transgenic (n=4)	2.605	0.217198066		1.588	4.6

¹Fresh fecal samples were collected on 3 different days was freeze-dried and then dried to constant weight at 110°C for 24 h, and analyzed for total phosphorus.

²At the 5% level of confidence t=2.57.

Table 11. Phytase activities of the first generation (G1) transgenic offspring obtained by the crossing the phytase positive boar 167-02 with non-transgenic Yorkshire gilts¹

ID # of pig	Birth Date	Sex	Salivary phytase (U/ml)	Specific Activity U/mg protein
151-01	Mar 16/00	F	1193	126
151-02	"	F	736	63.3
151-05	"	M	710	109
151-07	"	M	8019	315
152-04	"	M	10077	364
152-09	"	M	3054	200
154-01	Mar 19/00	F	2472	256
154-03	"	F	6425	706
154-04	"	F	n.d.	n.d.
154-05	"	M	2767	213
154-06	"	M	341	39
154-07	"	M	4029	142
154-08	"	M	1184	47.4
159-03	Mar 20/00	F	1563	116
159-04	"	M	2285	201

¹The number of males and females (M/F) in each litter were 5/3, 7/2, 5/4, and 2/3 for litter numbers 151, 152, 154 and 159, respectively. Saliva was collected from the piglets on day 11.

Table 12. Primers used for construction and detection of transgenic constructs.

Name	Start-End ¹	Forward/ Reverse	
Primers used in R15/APPA+intron and R15/APPA construction			
APPA-DOWN2		R	TCGGCGCTCACCTTGAGTC
APPA-DRA		F	<u>CCGTTAAAGCCATCTTAATCCCAT</u>
APPA-SMA		R	<u>GTCGGGGTATGCGTGCTTCATTC</u>
CAT-ATG		R	<u>CCATGGTGGCGGCTTTAGCTTCCTAGCT</u> CCTGA
CAT-TAA		F	<u>AGCGCTTGAGTTGTAAGGCAGTTATTG</u> GTGCC
CAT-UP1		F	TCG AGG AGC TTG GCG AGA TT
R15-UP1		F	TTTCGGGCCAATGTTGCTGT
Primers used in SV40/APPA+intron construction			
SV-HIND		F	CCCAAGCTTACACTTATGC
SV-XHO		R	GCCCTCGAGCCTCCTCACTACTCT
Primers used in Lama2/APPA and Lama2/PSP/APPA construction			
APPA-CLA	12635-12657	F	GGATCGATAAAAGCCGCCACCATGAA
APPA-DOWN2	13307-13326	R	TCGGCGCTCACCTTGAGTC
APPA-DOWN4	12751-12780	R	GCACGCACACCATGACGACTGACAATCAC C
APPA-KPN	13935-13959	R	<u>CGGGTACCTTACAAACTGCAAGCGG</u>
APPA-MATURE	12719-12738	F	CAGAGTGAGCCGGAGCTGAA
APPA-UP2	13210-13229	F	CGAACTGGAACGGGTGCTTA
LAMA-CLA	12615-12639	R	GCATCGATCTTGGTTCTGACAAATGG
LAMA-SIGNAL		R	TGACTCTGAGTTCCCAATGA
LAMA-UP	12111-12130	F	GTGCTGCTCCAAGTTGGTG
Primers for detection of the porcine β-globin gene			
PIG-BGF		F	GCAGATTCCAAACCTCGCAGAG
PIG-BGR		R	CTGCCCAAGTCCTAAATGTGCGT

1 The location of the primers shown for Lama2/APPA sequence.

The start and stop codons of APPA are indicated in bold letters, the optimal initiation sequence for translation is italicized, and the restriction sites for restriction enzymes are underlined.

Reference List

The following references have been referred to in the present application. The content of these references are incorporated herein by reference.

1. Abelson,P.H. 1999. A potential phosphate crisis [editorial]. *Science* **283**: 2015.
2. Altschul,S.F., Gish,W., Miller,W., Myers,E.W., and Lipman,D.J. 1990. Basic local alignment search tool. *J.Mol.Biol.* **215**: 403-410.
3. Ann,D.K. and Carlson,D.M. 1985. The structure and organization of a proline-rich protein gene of a mouse multigene family. *J.Biol.Chem.* **260**: 15863-15872.
4. Arkhipovets,A.I. 1956. Age characteristics of salivary secretion in young pigs. *Journal of Physiology of USSR* **42**: 882-886.
5. Association of Official Analytical Chemists (AOAC) 1984. Official methods of analysis of the Association of Official Analytical Chemists. Association of Official Analytical Chemists, Washington.
6. Ausbel,F.M., Brent,R., Kingston,R.E., Moore,D.D., Seidman,J.G., Smith,J.A., and Struhl,K. 1992. Short protocols in molecular biology. John Wiley & Sons, New York.
7. Barsh,G.S. and Epstein,C.J. 1989. Physical and genetic characterization of a 75-kilobase deletion associated with al, a recessive lethal allele at the mouse agouti locus. *Genetics* **121**: 811-818.
8. Bennick,A. 1982. Salivary proline-rich proteins. *Mol.Cell Biochem.* **45**: 83-99.
9. Bitar,K. and Reinhold,J.G. 1972. Phytase and alkaline phosphatase activities in intestinal mucosae of rat, chicken, calf, and man. *Biochim.Biophys.Acta* **268**: 442-452.
10. Burge,C. and Karlin,S. 1997. Prediction of complete gene structures in human genomic DNA. *J.Mol.Biol* **268**: 78-94.

11. Burge,C.B. and Karlin,S. 1998. Finding the genes in genomic DNA. *Curr.Opin.Struct.Biol.* 8: 346-354.
12. Carlson,D.M. 1993. Salivary proline-rich proteins: biochemistry, molecular biology, and regulation of expression. *Crit Rev.Oral Biol.Med.* 4: 495-502.
13. Carswell,S. and Alwine,J.C. 1989. Efficiency of utilization of the simian virus 40 late polyadenylation site: effects of upstream sequences. *Mol.Cell Biol.* 9: 4248-4258.
14. Chi,T.H. and Crabtree,G.R. 2000. Perspectives: signal transduction. Inositol phosphates in the nucleus [comment]. *Science* 287: 1937-1939.
15. Clark,A.J., Archibald,A.L., McClenaghan,M., Simons,J.P., Wallace,R., and Whitelaw,C.B. 1993. Enhancing the efficiency of transgene expression. *Philos.Trans.R.Soc.Lond B Biol.Sci.* 339: 225-232.
16. Clements,S., Mehansho,H., and Carlson,D.M. 1985. Novel multigene families encoding highly repetitive peptide sequences. Sequence analyses of rat and mouse proline-rich protein cDNAs. *J Biol.Chem.* 260: 13471-13477.
17. Corring,T. 1980. Endogenous secretions in the pig. In *Current concepts of digestion and absorption in pigs*. Edited by A.G.Low and I.G.Partridge. National Institute for Research in Dairying, Reading pp. 136-150.
18. Cosgrove,D.J. 1980. Inositol phosphates. Their chemistry, biochemistry and physiology. Elsevier, Amsterdam.
19. Crameri,A., Dawes,G., Rodriguez,E.Jr., Silver,S. and Stemmer,W.P.C. 1997. Molecular evolution of an arsenate detoxification pathway by DNA shuffling. *Nature Biotechnol.* 15: 436-438.
20. Dassa,J., Marck,C., and Boquet,P.L. 1990. The complete nucleotide sequence of the *Escherichia coli* gene *APPA* reveals significant homology between pH 2.5 acid phosphatase and glucose-1-phosphatase. *J.Bacteriol.* 172: 5497-5500.

21. Denny,H.R. and Messervy,A. 1972. Surgical techniques for the extirpation of the submandibular salivary glands and the collection of salivary secretions in the pig. *Vet.Rec.* 90: 650-654.
22. Dobrinsky,J.R., Johnson,L.A., and Rath,D. 1996. Development of a culture medium (BECM-3) for porcine embryos: Effect of bovine serum albumin and fetal bovine serum on embryo development. *Biol.Reprod.* 55: 1069-1074.
23. Dvorakova,J. 1998. Phytase: sources, preparation and exploitation. *Folia Microbiol.(Praha)* 43: 323-338.
24. Engelen,A.J., van der Heeft,F.C., Randsdorp,P.H., and Smit,E.L. 1994. Simple and rapid determination of phytase activity. *J AOAC Int.* 77: 760-764.
25. Frandson,R.D. and Spurgeon,T.L. 1992. Anatomy and physiology of farm animals. *Lea & Febiger, Philadelphia.*
26. Galfre,G. and Milstein,C. 1981. Preparation of monoclonal antibodies: strategies and procedures. *Methods Enzymol.* 73: 3-46.
27. Geyer,P.K. 1997. The role of insulator elements in defining domains of gene expression. *Curr.Opin.Genet.Dev.* 7: 242-248.
28. Gish,W. and States,D.J. 1993. Identification of protein coding regions by database similarity search. *Nature Genetics* 3: 266-272.
29. Golovan,S., Wang,G., Zhang,J., and Forsberg,C.W. 2000. Characterization and overproduction of the *Escherichia coli* APPA encoded bifunctional enzyme that exhibits both phytase and acid phosphatase activities [In Process Citation]. *Can.J.Microbiol.* 46: 59-71.
30. Gordon,J.W., Scangos,G.A., Plotkin,D.J., Barbosa,J.A., and Ruddle,F.H. 1980. Genetic transformation of mouse embryos by microinjection of purified DNA. *Proc.Natl.Acad.Sci.USA* 77: 7380-7384.

31. Gorman,C., Padmanabhan,R., and Howard,B.H. 1983. High efficiency DNA-mediated transformation of primate cells. *Science* 221: 551-553.
32. Greiner, R. and Jany, K.-D. Characterization of a phytase from *Escherichia coli*. Herbsttagung der Gesellschaft fur Biologische Chemie 372. 1991.
Ref Type: Abstract
33. Greiner,R., Konietzny,U., and Jany,K.D. 1993. Purification and characterization of two phytases from *Escherichia coli*. *Arch Biochem.Biophys.* 303: 107-113.
34. Hall,J., Hazlewood,G.P., Surani,M.A., Hirst,B.H., and Gilbert,H.J. 1990. Eukaryotic and prokaryotic signal peptides direct secretion of a bacterial endoglucanase by mammalian cells. *J.Biol.Chem.* 275: 19996-19999.
35. Harland,B.F. and Morris,E.R. 1995. Phytate: A good or a bad food component? *Nutr.Res.* 15: 733-754.
36. Harlow,E., and Lane, D. 1988. *Antibodies: a laboratory manual*. Cold Spring Harbor Laboratory, New York.
37. Harayama,S. 1998. Artificial evolution by DNA shuffling. *Trends Biotechnol.* 16: 76-81.
38. Heinoen,J.K. and Lahti,R.J. 1981. A new and convenient colorimetric determination of inorganic orthophosphate and its application to the assay of inorganic pyrophosphate. *Anal.Biochem.* 133: 313-317.
39. Heneine,W. and Switzer,W.M. 1996. Highly sensitive and specific polymerase chain reaction assays for detection of baboon and pig cells following xenotransplantation in humans. *Transplantation* 62: 1360-1362.
40. Higuchi,R. 1989. Simple and rapid preparation of samples for PCR. In *PCR technology : principles and applications for DNA amplification*. Edited by H.A.Erlich. Macmillan Publishers ; Stockton Press, New York, NY pp. 31-38.

41. Hu,Y., Nakagawa,Y., Purushotham,K.R., and Humphreys-Beher,M.G. 1992. Functional changes in salivary glands of autoimmune disease-prone NOD mice. *Am.J.Physiol.* 263: E607-E614.
42. Huang,X. 1996. An improved sequence assembly program. *Genomics* 33: 21-31.
43. Huang,X. 1999. A contig assembly program based on sensitive detection of fragment overlaps. *Genomics* 14: 18-25.
44. Huff,W.E., Moore,P.A., Jr., Waldroup,P.W., Waldroup,A.L., Balog,J.M., Huff,G.R., Rath,N.C., Daniel,T.C., and Raboy,V. 1998. Effect of dietary phytase and high available phosphorus corn on broiler chicken performance. *Poult Sci* 77: 1899-1904.
45. Iqbal,T.H., Lewis,K.O., and Cooper,B.T. 1994. Phytase activity in the human and rat small intestine. *Gut* 35: 1233-1236.
46. Jia,Z., Golovan,S., Ye,Q., and Forsberg,C.W. 1998. Purification, crystallization and preliminary X-ray analysis of the *Escherichia coli* phytase. *Acta Crystallogr.D.Biol.Crystallogr.* 54 (Pt 4): 647-649.
47. Kazazian,H.H.J. 1998. Mobile elements and disease. *Curr.Opin.Genet.Dev.* 8: 343-350.
48. Kim,H.S. and Maeda,N. 1986. Structures of two HaeIII-type genes in the human salivary proline-rich protein multigene family. *J Biol.Chem.* 261: 6712-6718.
49. Kohler, G. and Milstein, C. 1976. Derivation of specific antibody-producing tissue culture and tumor lines by cell fusion. *Eur. J. Immunol.* 7: 511-519.
50. Kozak,M. 1987. At least six nucleotides preceding the AUG initiator codon enhance translation in mammalian cells. *Proc.Natl.Acad.Sci.USA* 196: 947-950.
51. Kuchner,O. and Arnold,F.H. 1997. Directed evolution of enzyme catalysts. *Trends Biotechnol.* 15: 523-530.
52. Laemmli,U.K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227: 680-785.

53. Larsen,H.J., Brodersen,C.H., and Hjorth,J.P. 1994. High-level salivary gland expression in transgenic mice. *Transgenic Res.* 3: 311-316.
54. Laursen,J. and Hjorth,J.P. 1997. A cassette for high-level expression in the mouse salivary glands. *Gene* 198: 367-372.
55. Levitt,N., Briggs,D., Gil,A., and Proudfoot,N.J. 1989. Definition of an efficient synthetic poly(A) site. *Genes Dev.* 3: 1019-1025.
56. Low,A.G. 1989. Research into the digestive physiology of pigs. *In Nutrition and digestive physiology in monogastric farm animals. Edited by E.J.van Weerden and J.Huisman. Pudoc, Wageningen pp. 1-15.*
57. Madsen,H.O. and Hjorth,J.P. 1985. Molecular cloning of mouse PSP mRNA. *Nucleic Acids Res.* 13: 1-13.
58. Mallin,M.A. 2000. Impacts of industrial animal production on rivers and estuaries. *American Scientist* 88: 26-37.
59. Mehansho,H., Ann,D.K., Butler,L.G., Rogler,J., and Carlson,D.M. 1987. Induction of proline-rich proteins in hamster salivary glands by isoproterenol treatment and an unusual growth inhibition by tannins. *J.Biol.Chem.* 262: 12344-12350.
60. Menniti,F.S., Oliver,K.G., Putney,J.W., Jr., and Shears,S.B. 1993. Inositol phosphates and cell signaling: new views of InsP5 and InsP6. *Trends Biochem.Sci* 18: 53-56.
61. Mikkelsen,T.R., Brandt,J., Larsen,H.J., Larsen,B.B., Poulsen,K., Ingerslev,J., Din,N., and Hjorth,J.P. 1992. Tissue-specific expression in the salivary glands of transgenic mice. *Nucleic Acids Res.* 20: 2249-2255.
62. Miner,J.R. 1999. Alternatives to minimize the environmental impact of large swine production units. *J.Anim Sci.* 77: 440-444.
63. Nakai,K. and Kanehisa,M. 1992. A knowledge base for predicting protein localization sites in eukaryotic cells. *Genomics* 14: 897-911.

64. Nesterenko,M.V., Tilley,M., and Upton,S.J. 1994. A simple modification of Blum's silver stain method allows for 30 minute detection of proteins in polyacrylamide gels. *J Biochem.Biophys.Methods* 28: 239-242.
65. O'Shannessy,D.J., Voorstad,P.J., and Quarles,R.H. 1987. Quantitation of glycoproteins on electroblots using the biotin- streptavidin complex. *Anal.Biochem.* 163: 204-209.
66. Owerbach,D. and Hjorth,J.P. 1980. Inheritance of a parotid secretory protein in mice and its use in determining salivary amylase quantitative variants. *Genetics* 95: 129-141.
67. Pen,J., Verwoerd,T.C., Vanparidon,P.A., Beudeker,R.F., Vandenelzen,P.J.M., Geerse,K., Vanderklis,J.D., Versteegh,H.A.J., Vanooyen,A.J.J., and Hoekema,A. 1993. Phytase-containing transgenic seeds as a novel feed additive for improved phosphorus utilization. *Biotechnology* 11: 811-814.
68. Petitclerc,D., Attal,J., Theron,M.C., Bearzotti,M., Bolifraud,P., Kann,G., Stinnakre,M.G., Pointu,H., Puissant,C., and Houdebine,L.M. 1995. The effect of various introns and transcription terminators on the efficiency of expression vectors in various cultured cell lines and in the mammary gland of transgenic mice. *J Biotechnol.* 40: 169-178.
69. Pinkert,C.A., Dyer,T.J., Kooyman,D.L., and Kiehm,D.J. 1990. Characterization of transgenic livestock production. *Domest.Anim Endocrinol.* 7: 1-18.
70. Quissell,D.O., Barzen,K.A., Redman,R.S., Camden,J.M., and Turner,J.T. 1998. Development and characterization of SV40 immortalized rat parotid acinar cell lines. *In Vitro Cell Dev.Biol.-Animal.* 34: 58-67.
71. Reddy,N.R., Sathe,S.K., and Salunkhe,D.K. 1982. Phytates in legumes and cereals. *Adv.Food Res.* 28: 1-92.
72. Rozhkov,Y.I. and Galimov,I.R. 1990. Salivary gland polymorphism in pigs and cattle detected by affinity electrophoresis. *Anim.Genet* 21: 277-283.

73. Samuelson,L.C. 1996. Transgenic approaches to salivary gland research. *Annu.Rev.Physiol.* 58: 209-229.
74. Shaw,P. and Schibler,U. 1986. Structure and expression of the parotid secretory protein gene of mouse. *J Mol.Biol.* 192: 567-576.
75. Simons,P.C., Versteegh,H.A., Jongbloed,A.W., Kemme,P.A., Slump,P., Bos,K.D., Wolters,M.G., Beudeker,R.F., and Verschoor,G.J. 1990. Improvement of phosphorus availability by microbial phytase in broilers and pigs. *Br J Nutr.* 64: 525-540.
76. Smit,A.F.A. 1996. The origin of interspersed repeats in the human genome. *Curr.Opin.Genet.Dev.* 6: 743-748.
77. Stahl,C.H., Han,Y.M., Roneker,K.R., House,W.A., and Lei,X.G. 1999. Phytase improves iron bioavailability for hemoglobin synthesis in young pigs. *J Anim Sci* 77: 2135-2142.
78. Svendsen,P., Laursen,J., Krogh-Pedersen,H., and Hjorth,J.P. 1998a. Novel salivary gland specific binding elements located in the PSP proximal enhancer core. *Nucleic Acids Res.* 26: 2761-2770.
79. Taboit-Dameron,F., Malassagne,B., Viglietta,C., Puissant,C., Leroux-Coyau,M., Chereau,C., Attal,J., Weill,B., and Houdebine,L.M. 1999. Association of the 5'HS4 sequence of the chicken beta-globin locus control region with human EF1 alpha gene promoter induces ubiquitous and high expression of human CD55 and CD59 cDNAs in transgenic rabbits. *Transgenic Res.* 8: 223-235.
80. Terada,T. and Nakamura,Y. 1993. An immunohistochemical survey of amylase isoenzymes in cholangiocarcinoma and hepatocellular carcinoma. *Arch Pathol.Lab Med.* 117: 160-162.
81. Towbin, H., Staehelin, T., and Gordon, J. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc.Natl.Acad.Sci.USA* 76(9), 4350-4354. 1979.

Ref Type: Generic

82. Tryon,A.F. and Bibby,B.G. 1966. Preliminary studies on pig saliva. *Arch.Oral Biol.* 11: 527-531.
83. Tu,Z.J., Lazowski,K.W., Ehlenfeldt,R.G., Wu,G., Lin,H.H., Kousvelari,E., and Ann,D.K. 1993. Isoproterenol/tannin-dependent R15 expression in transgenic mice is mediated by an upstream parotid control region. *Gene Expr.* 3: 289-305.
84. Verwoerd,T.C., van Pardon,P.A., van Ooyen,A.J., van Lent,J.W., Hoekema,A., and Pen,J. 1995. Stable accumulation of *Aspergillus niger* phytase in transgenic tobacco leaves. *Plant Physiol* 109: 1199-1205.
85. Wall,R.J., Hyman,P., Kerr,D., Pintado,B., and Wells,K. 1997. Transgenic animal technology. *J.Androl* 18: 236-239.
86. Wall,R.J., Pursel,V.G., Hammer,R.E., and Prinster,R.L. (*Editors*). 1985. Development of porcine ova that were centrifuged to permit visualization of pronuclei and nuclei. *Biol.Reprod.* 32.
87. Wessel, D. and Flugge, U.I. 1984. A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids. *Anal.Biochem.* 138: 141-143.
88. Whitelaw,C.B., Archibald,A.L., Harris,S., McClenaghan,M., Simons,J.P., and Clark,A.J. 1991. Targeting expression to the mammary gland: intronic sequences can enhance the efficiency of gene expression in transgenic mice. *Transgenic Res.* 1: 3-13.
89. Williamson,C.M., A.J.Bramley, and A.J.Lax 1994. Expression of the lysostaphin gene of *Staphylococcus simulans* in a eukaryotic system. *Appl.Environ.Microbiol.* 60: 771-776.
90. Wodzinski,R.J. and Ullah,A.H. 1996. Phytase. *Adv.Appl.Microbiol.* 42: 263-302.
91. Xiong,Y. and Eickbush,T.H. 1990. Origin and evolution of retroelements based upon their reverse transcriptase sequences. *EMBO J.* 9: 3353-3362.

THE EMBODIMENTS OF THE INVENTION IN WHICH AN EXCLUSIVE PROPERTY OR PRIVILEGE IS CLAIMED ARE DEFINED AS FOLLOWS:

1. A transgenic non-human animal that carries in the genome of its somatic and/or germ cells a nucleic acid sequence including a heterologous transgene construct, said construct including a transgene encoding a protein, said transgene being operably linked to a first regulatory sequence for salivary gland specific expression of said protein.
2. The animal of claim 1 wherein said first regulatory sequence comprises a saliva protein promoter/enhancer sequence, whereby said animal expresses said protein in its saliva.
3. The animal of claim 1 wherein said animal is a mammal.
4. The animal of claim 3 wherein said animal is chosen from the group comprising pigs, goats, sheep, cows, horses, rabbits, rodents, cats and dogs, and in addition, fish and poultry.
5. The animal of claim 1 wherein said saliva protein promoter/enhancer sequence comprises a parotid secretory protein (PSP) promoter/enhancer, a proline-rich protein (PRP) promoter/enhancer or a salivary amylase promoter/enhancer.
6. The animal of claim 5 wherein said promoter/enhancer is a parotid secretory protein (PSP) promoter/enhancer.
7. The animal of claim 6 wherein said parotid secretory protein (PSP) promoter/enhancer is derived from a mouse.
8. The animal of claim 5 wherein said promoter/enhancer is a proline-rich protein (PRP) promoter/enhancer.
9. The animal of claim 8 wherein said proline-rich protein (PRP) promoter/enhancer is derived from a rat.

10. The animal of claim 1 wherein said transgene is further operably linked to one or more second regulatory sequences including enhancers, transcription regulatory sequences, termination sequences, and polyadenylation sites.

5 11. The animal of claim 1 wherein said transgene comprises a gene encoding a protein having phytase activity.

12. The animal of claim 1 wherein said transgene encodes a phytase or a homologue thereof.

10 13. The animal of claim 1 wherein said animal is a pig, said transgene comprising a gene encoding a protein having phytase activity and wherein said first regulatory sequence comprises a parotid secretory protein (PSP) promoter/enhancer or a proline-rich protein (PRP) promoter/enhancer.

15 14. The animal of claim 1 wherein said transgene construct comprises a nucleic acid sequence according to SEQ ID NO:3, SEQ ID NO:5; or SEQ ID NO:7.

20 15. A transgenic non-human animal that carries in the genome of its somatic and/or germ cells a nucleic acid sequence including a heterologous transgene construct, said construct including a transgene encoding phytase or a homologue thereof.

16. The animal of claim 15 wherein said transgene is operably linked to a first regulatory sequence for salivary gland specific expression of said phytase.

25 17. The animal of claim 16 wherein said first regulatory sequence comprises a parotid secretory protein (PSP) promoter/enhancer, a proline-rich protein (PRP) promoter/enhancer or a salivary amylase promoter/enhancer.

30 18. The animal of claim 17 wherein said animal is a mammal.

19. The animal of claim 18 wherein said phytase or a homologue thereof is expressed in saliva or in the gastrointestinal tract of said animal.
20. The animal of claim 15 wherein said transgene construct comprises a nucleic acid sequence according to SEQ ID NO:3, SEQ ID NO:5; or SEQ ID NO:7.
21. A method of expressing a protein, the method comprising the steps of:
 - a) introducing a transgene construct into a non-human animal embryo such that a non-human transgenic animal that develops from said embryo has a genome that comprises said transgene construct, wherein said transgene construct comprises:
 - i) a transgene encoding said protein, and
 - ii) at least one regulatory sequence for gastrointestinal tract specific expression of said protein,
 - b) transferring said embryo to a foster female; and,
 - c) developing said embryo into said transgenic animalwherein said transgene is produced in the gastrointestinal tract of said animal.
22. The method of claim 21 wherein said regulatory sequence provides for salivary gland or pancreatic gland specific expression of said protein.
23. The method of claim 21 wherein said regulatory sequence provides for salivary gland specific expression of said protein.
24. The method of claim 23 wherein said salivary gland is a parotid gland, submaxillary gland, or a submandibular gland.
25. The method of claim 23 wherein said transgene is expressed in the saliva of said animal.
- 30 26. The method of claim 21 wherein said transgene is heterologous.

27. The method of claim 21 wherein said at least one regulatory sequence comprises a salivary protein promoter/enhancer sequence.

28. The method of claim 21 wherein said protein is a glycoprotein.

5

29. A transgenic animal adapted for expressing a protein according to the method of claim 21, or a progeny thereof.

10

30. The method of claim 21 wherein said protein is a phytase or a homologue thereof.

15

31. The method of claim 21 wherein said transgene construct comprises a nucleic acid sequence according to SEQ ID NO:3, SEQ ID NO:5, or SEQ ID NO:7.

20

32. A process for producing a protein comprising the steps of:

15 a) obtaining saliva containing said protein from a non-human transgenic animal, said animal containing within its genome a transgene construct, wherein said transgene construct comprises:

20 i) a transgene encoding said protein, and
ii) at least one regulatory sequence for salivary gland specific expression of
said protein, and
extracting said protein from said saliva.

33. The process of claim 32 wherein said transgene is heterologous.

25

34. The process of claim 32 wherein said at least one regulatory sequence comprises a salivary protein promoter/enhancer sequence.

30

35. The process of claim 32 wherein said protein is a glycoprotein.

36. The process of claim 32 wherein said transgene construct comprises a nucleic acid sequence according to SEQ ID NO:3, SEQ ID NO:5; or SEQ ID NO:7.

37. The process of claim 32 wherein said protein is a phytase or a homologue thereof.

38. The process of claim 32 wherein said salivary gland is a parotid gland, submaxillary, or a submandibular gland.

5

39. A method for expressing a phytase or a homologue thereof in a non-human animal, said method comprising:

a) constructing a nucleic acid sequence including a transgene construct comprising:

i) a transgene encoding said phytase or a homologue thereof, and

10 ii) at least one regulatory sequence for gastrointestinal tract specific expression of said protein, and

b) transfected the animal with said nucleic acid sequence;

whereby said animal carries within the genome of its somatic and/or germ cells said transgene construct and wherein said animal expresses said phytase or a homologue thereof 15 in its gastrointestinal tract.

40. The method of claim 39 wherein said transgene construct results in salivary gland or pancreatic gland specific expression of said phytase or a homologue thereof.

20 41. The method of claim 40 wherein said regulatory sequence provides for salivary gland specific expression of said phytase or a homologue thereof.

42. The method of claim 41 wherein said salivary gland is a parotid gland, submaxillary, or a submandibular gland.

25

43. The method of claim 41 wherein said phytase or a homologue thereof is expressed in the saliva of said mammal.

44. The method of claim 41 wherein said transgene construct comprises a nucleic acid 30 sequence according to SEQ ID NO:3, SEQ ID NO:5; or SEQ ID NO:7.

45. The method of claim 39 wherein said nucleic acid sequence is introduced into said animal in the form of a transgene construct.

46. The method of claim 45 wherein said transgene construct is a nucleic acid molecule.

5

47. The method of claim 46 wherein said plasmid comprises a nucleic acid sequence according to SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, or SEQ ID NO:6.

48. The method of claim 39 wherein said animal is chosen from the group comprising pigs, goats, sheep, cows, horses, rabbits, rodents, cats, dogs, fish and poultry.

10 49. The method of claim 48 wherein said animal comprises a mouse or a pig.

50. A nucleic acid molecule comprising a nucleic acid sequence including a gene 15 encoding a protein, said gene being operably linked to at least one regulatory sequence for gastrointestinal tract specific expression of said protein.

51. The molecule of claim 50 wherein said at least one regulatory sequence comprises a salivary protein promoter/enhancer sequence, whereby expression of said protein is salivary 20 gland specific.

52. The molecule of claim 51 wherein said salivary protein promoter/enhancer sequence comprises a parotid secretory protein (PSP) promoter/enhancer, a proline-rich protein (PRP) promoter/enhancer, a salivary amylase promoter/enhancer, or a SV40 promoter/enhancer.

25

53. The molecule of claim 51 wherein said protein comprises a phytase or a homologue thereof.

30 54. The molecule of claim 53 wherein said molecule is a transgene construct.

55. The molecule of claim 54 wherein said molecule is a nucleic acid molecule.

56. The molecule of claim 55 wherein said molecule comprises a nucleic acid sequence according to SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:4, or SEQ ID NO:6.

57. The molecule of claim 53 wherein said molecule includes a nucleic acid sequence according to SEQ ID NO:3, SEQ ID NO:5; or SEQ ID NO:7.

58. An antibody specific to a protein expressed by a nucleic acid sequence according to SEQ ID NO:3, SEQ ID NO:5; or SEQ ID NO:7.

10 59. The antibody of claim 58 wherein said antibody is monoclonal.

60. The antibody of claim 58 wherein said antibody is polyclonal.

61. A hybridoma secreting the antibody of claim 59.

15 62. A host cell transfected with molecule of claim 50.

63. A host cell transfected with the molecule of claim 56.

20 64. A host cell transfected with the molecule of claim 57.

65. The host cell of claim 63 wherein said cell is an bacterial cell.

66. The host cell of claim 64 wherein said cell is an animal cell.

25 67. A diagnostic kit for immunologically detecting a protein expressed by a nucleic acid sequence according to SEQ ID NO:3, SEQ ID NO:5; or SEQ ID NO:7, the kit including an antibody specific to said protein.

30 68. The kit of claim 67 wherein said antibody is monoclonal.

69. The kit of claim 68 wherein said antibody is polyclonal.

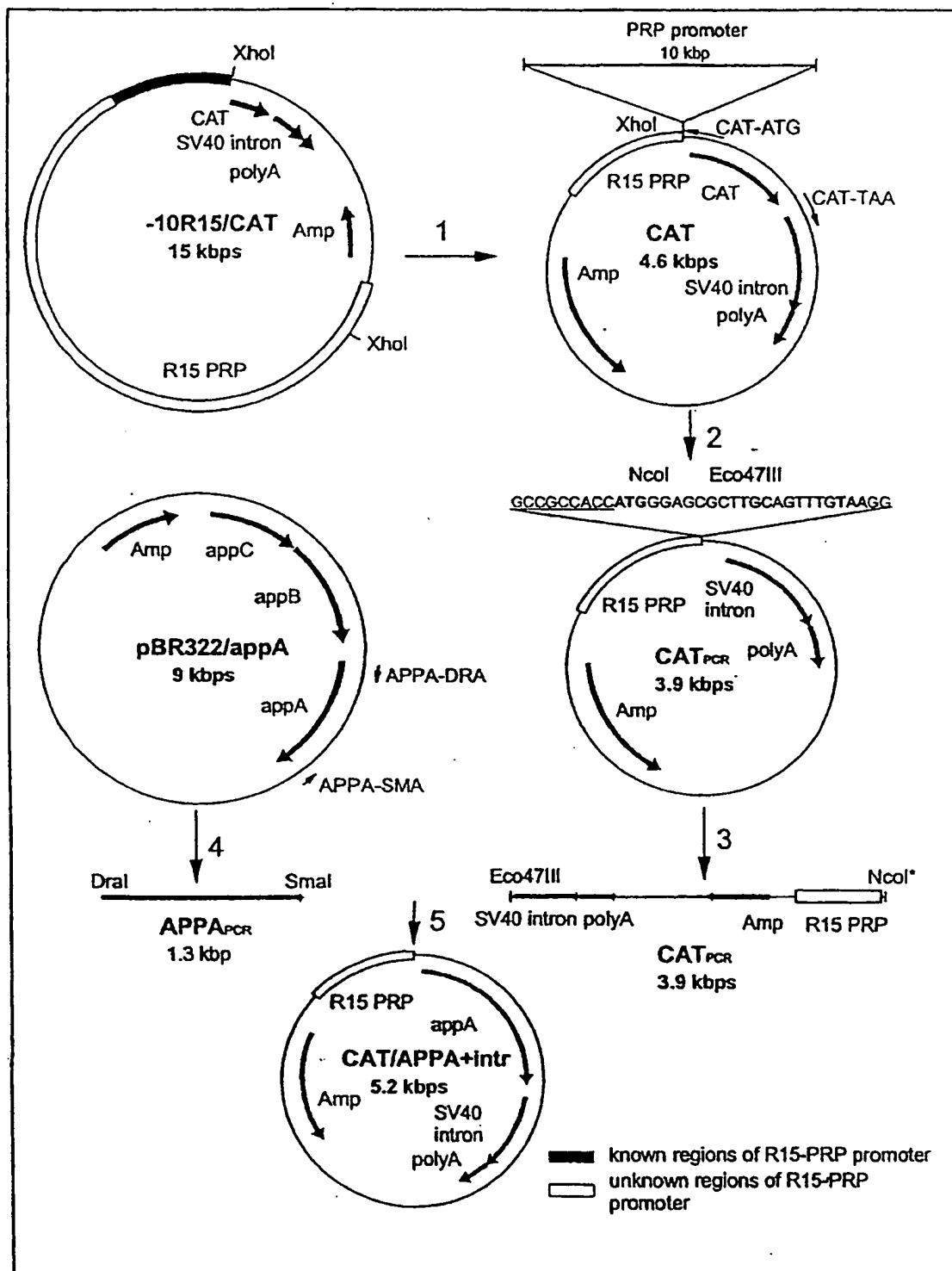
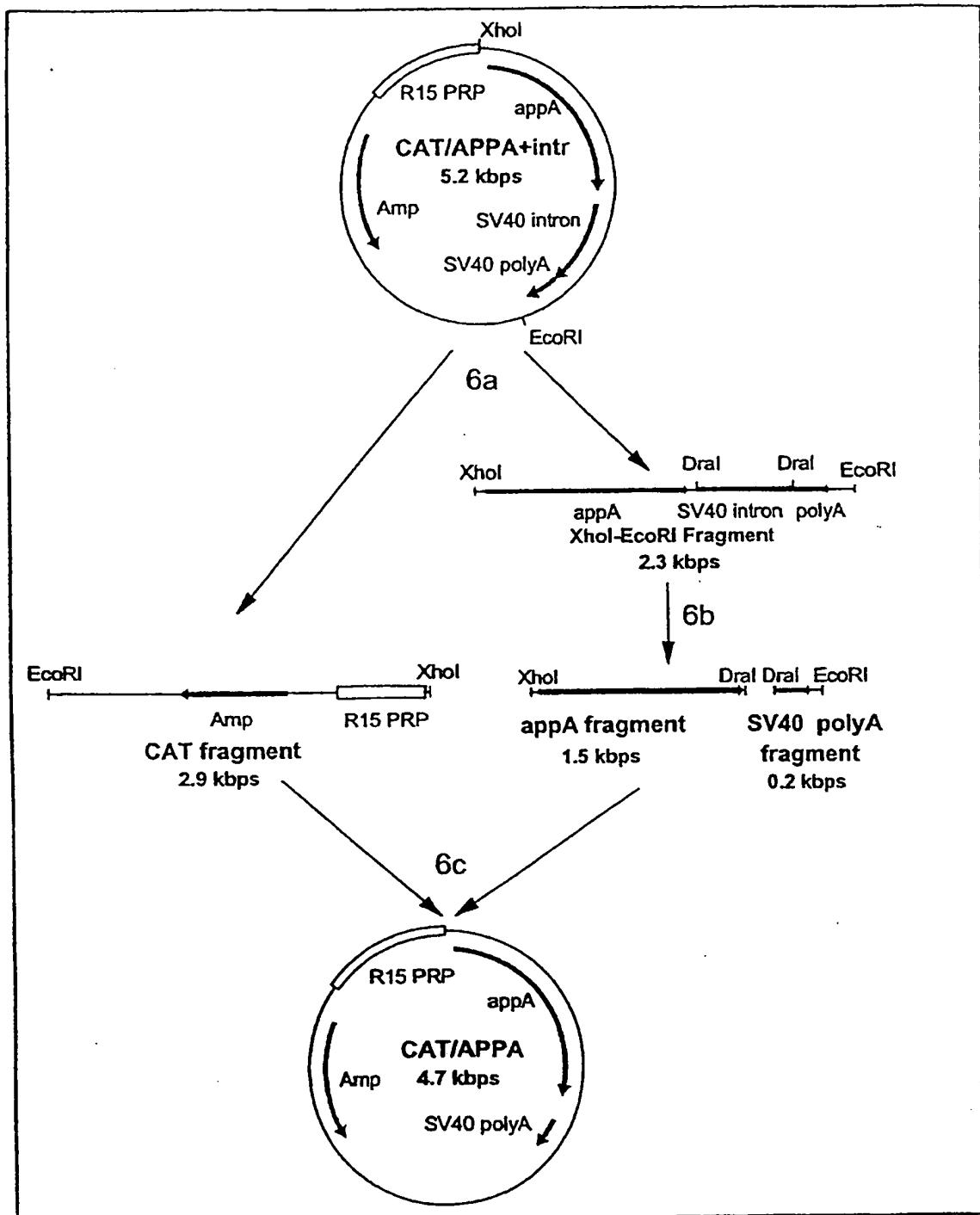
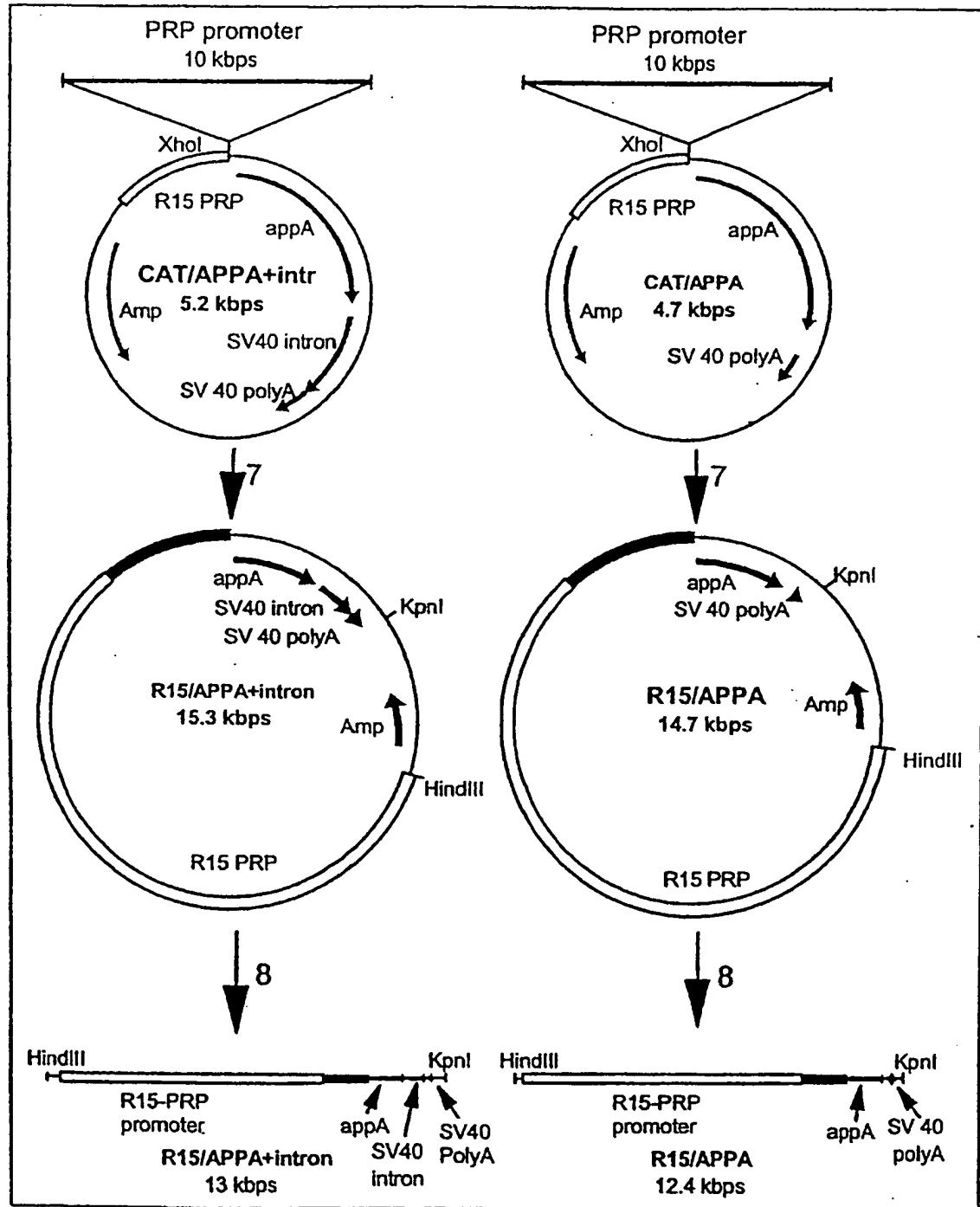
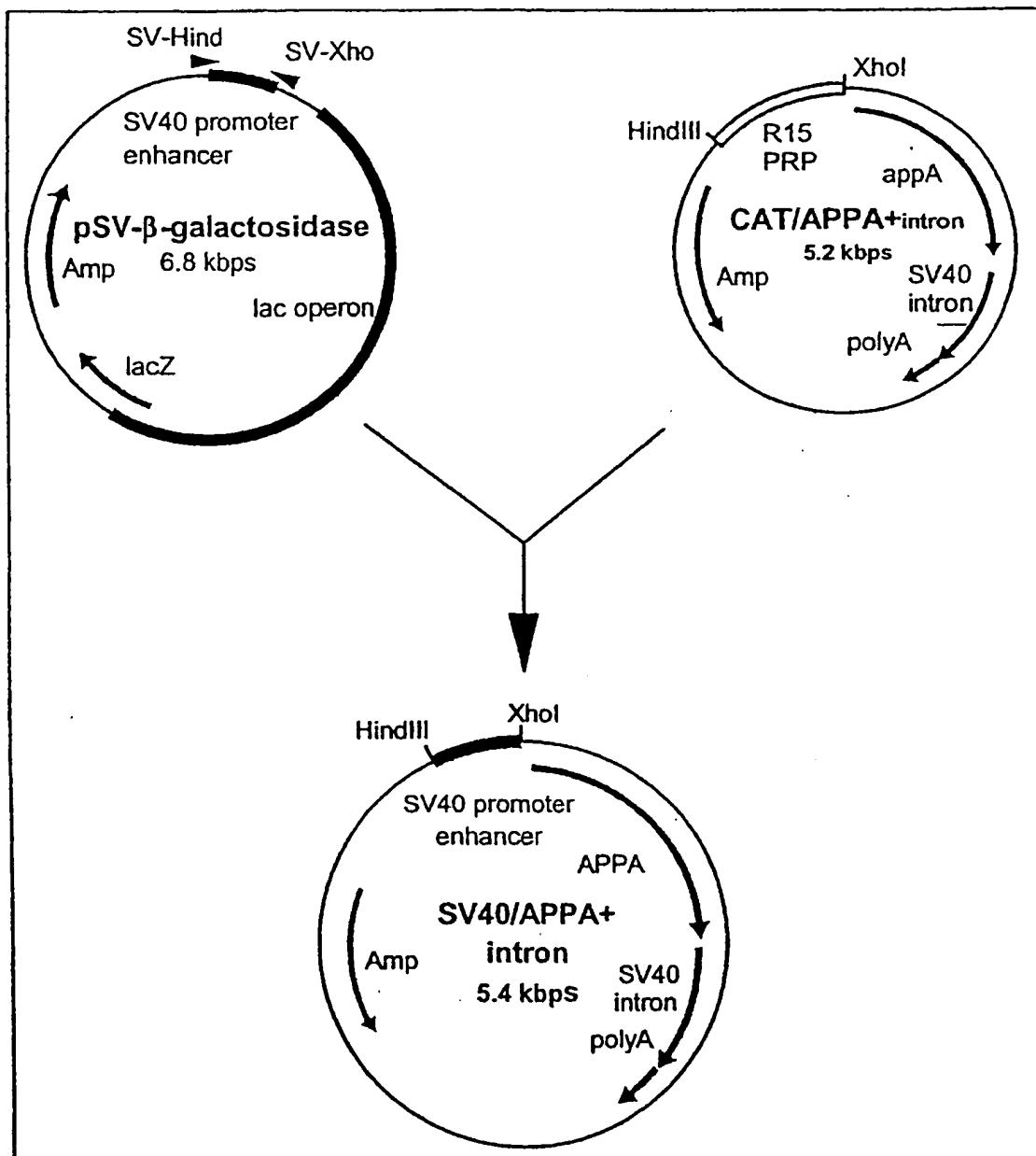
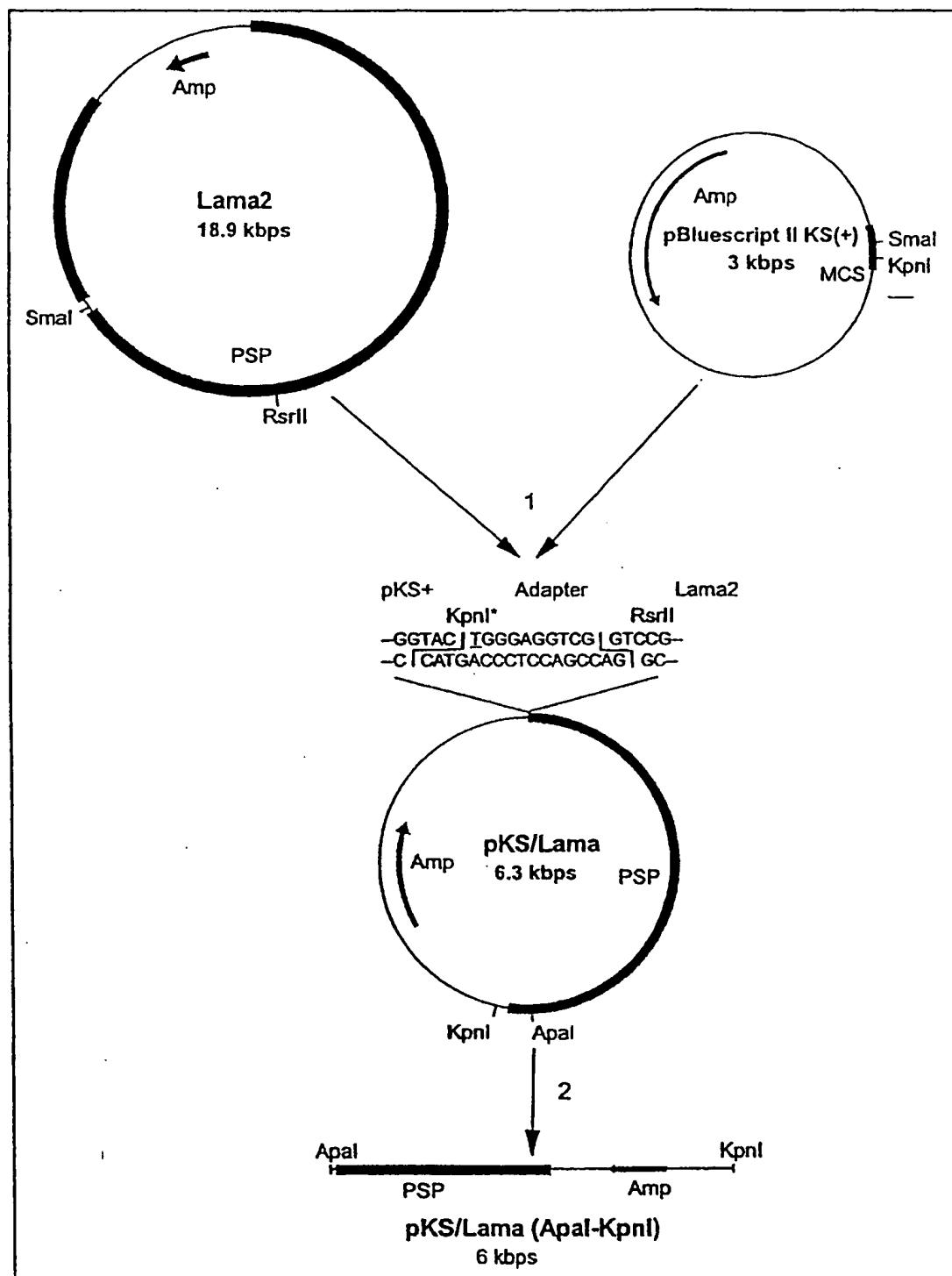


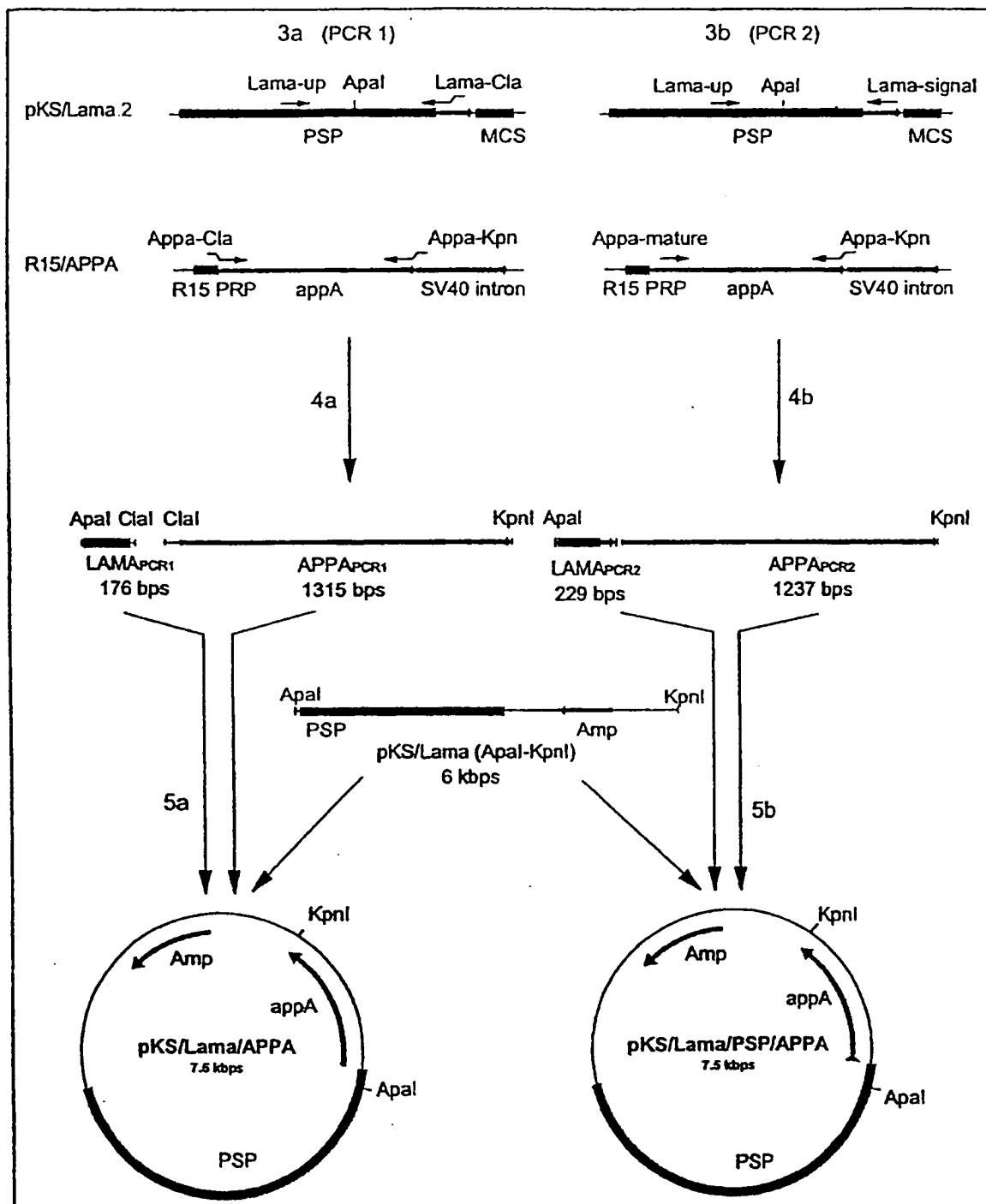
Figure 1

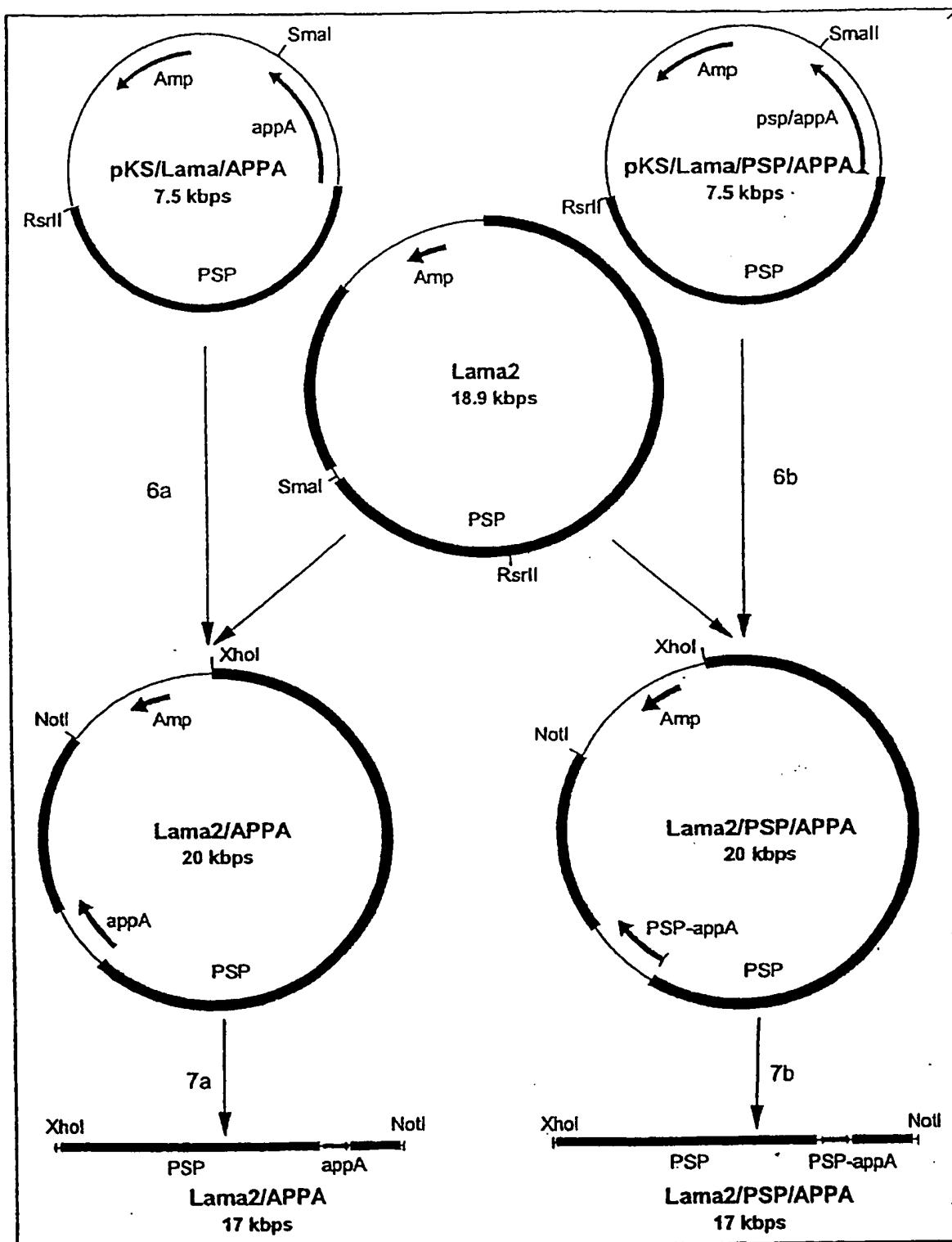
**Figure 1 (continued)**

**Figure 1 (continued)**

**Figure 2**

**Figure 3**

**Figure 3 (continued)**

**Figure 3 (continued)**

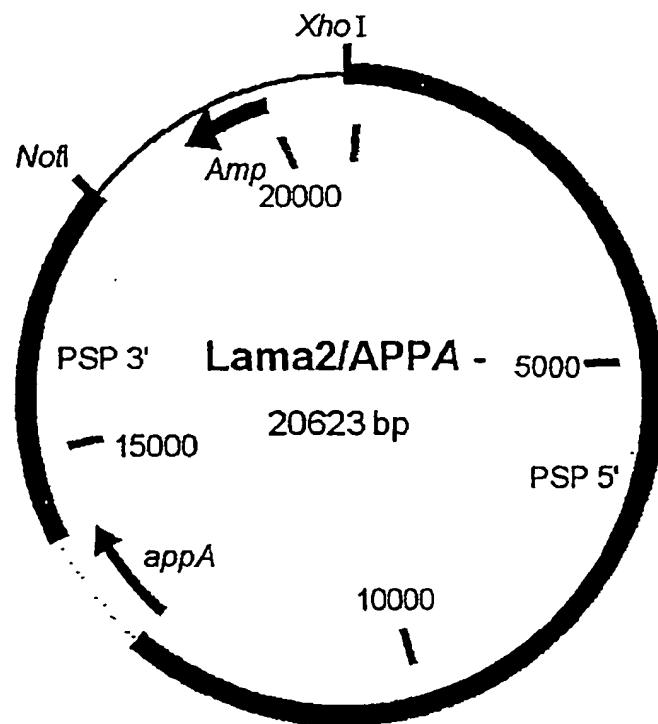


Figure 4. Schematic diagram of the Lama2/APPA construct.

Figure 5. The nucleic acid sequence of the Lama2/APPA plasmid (SEQ ID NO: 1)

LOCUS Lama-appA 20623 bp DNA CIRCULAR SYN 17-JAN-2000
DEFINITION Lama 2/APPA transgenic construct
ACCESSION Lama 2-appA,
KEYWORDS parotid secretory protein; acid glucose-1-phosphatase; appA
 gene;
 periplasmic phosphoanhydride phosphohydrolase; artificial
 sequence;
 cloning vector
REFERENCE 1 (bases 1 to 20623)
AUTHORS Golovan, S., Forsberg, C.W., Phillips, J.
JOURNAL Unpublished.
FEATURES
DEFINITION *M. musculus* Psp gene for parotid secretory protein.
ACCESSION X68699
VERSION X68699.1 GI:53809
SOURCE house mouse.
ORGANISM *Mus musculus*
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; *Mus*.
REFERENCE 1 (bases 3777 to 5332)
AUTHORS Svendsen, P., Laursen, J., Krogh-Pedersen, H. and Hjorth, J.P.
TITLE Novel salivary gland specific binding elements located in the PSP
 proximal enhancer core
JOURNAL Nucleic Acids Res. 26 (11), 2761-2770 (1998)
MEDLINE 98256451
REFERENCE 2 (bases 7147 to 12653; 13952 to 17731)
AUTHORS Mikkelsen, T.R.
TITLE Direct Submission
JOURNAL Submitted (07-OCT-1992) T.R. Mikkelsen, Department of Molecular
 Biology, University of Aarhus, CF Mollers Alle 130, 8000
 Aarhus, DENMARK
REFERENCE 3 (bases 7147 to 12653; 13952 to 17731)
AUTHORS Laursen, J., Hjorth, J.P.
TITLE A cassette for high-level expression in the mouse salivary glands.
JOURNAL Gene 1997 Oct 1;198(1-2):367-72
MEDLINE 9370303

FEATURES Location/Qualifiers
 source 1..to 12653; 13952 to 17731
 /organism="Mus musculus"
 /strain="C3H/As"
 /db_xref="taxon:10090"
 /chromosome="2"
 /map="Estimate: 69 cM from centromere"
 /clone="Lambda YP1, Lambda YP3, Lambda YP7"
 /clone_lib="Lambda-PHAGE (Lambda L47.1)"
 /germline
 /note="Allele: b"
 misc_feature 3777-5332
 /gene="PSP"
 /function="salivary gland specific positive acting
 regulatory region"
 enhancer 7147..8724
 /evidence=experimental
 exon 11778..11824
 /gene="Psp"
 /note="exon a"
 /number=1
 /evidence=experimental
 exon 12626.. 14190
 /gene="Psp"
 /note="exon b fused with exons h and i"
 misc_feature 12644-12652

Figure 5 (continued):

```

        /function=" consensus sequence for initiation in higher
        eukaryotes "
  misc_feature 13952-13965
        /function=" M13mp18 polylinker"

DEFINITION E. coli periplasmic phosphoanhydride phosphohydrolase (appA) gene,

  ACCESSION      M58708 L03370 L03371 L03372 L03373 L03374 L03375
  VERSION        M58708.1 GI:145283
  SOURCE         Escherichia coli DNA.
  ORGANISM        Escherichia coli
                  Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
                  Escherichia.

  REFERENCE 1 (bases 12653..13951)
  AUTHORS Dassa,J., Marck,C. and Boquet,P.L.
  TITLE  The complete nucleotide sequence of the Escherichia coli gene appA
         reveals significant homology between pH 2.5 acid phosphatase
         and glucose-1-phosphatase
  JOURNAL J. Bacteriol. 172 (9), 5497-5500 (1990)
  MEDLINE 90368616

  FEATURES          Location/Qualifiers
  Source            12653..13951
                    /organism="Escherichia coli"
                    /db_xref="taxon:562"
  sig_peptide       12653..12718
  /gene="appA"
  CDS12653          13951
                    /gene="appA"
                    /standard_name="acid phosphatase/phytase"
                    /transl_table=11
                    /product="periplasmic phosphoanhydride phosphohydrolase"
                    /protein_id="AAA72086.1"
                    /db_xref="GI:145285"

/translation="MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAP
TKATQLMQDVTDAWPTWPVKLGWLTPRGGEIAYLGHYORQRLVADGLLAKRGCPQS
GQVAILADVDERTRKTGEAAGLAPDCAITVHTQADTSSPDPFLNPLCTGVQQLDNA
NVTDAILSRAGGSIADFTGHRQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPS
ELKVSADNVSITGAVSIAASMLTEIFLLQQAQGMPEPGWGRITDSHEQWNTLSSLHNAQF
YLLQRTPEVARSRATPLLDLIKTAUTPHPPQKQAYGVTLPTSVLFIAGHDTNLNLANLGG
ALELNWTLPGQPDNTPPGGELVFERWRRLSDNSQWIQSVLVFQTLQQMRDKTPSLNT
PPGEVKLTLAGCEERNAQGMCSLAGFTQIVNEARIPACSL"
  mat_peptide       12719 13948
                    /gene="appA"
                    /product="periplasmic phosphoanhydride phosphohydrolase"

  mutation          replace(12659.. 12661,"gcg changed to gcc")
                    /gene="appA"
                    /standard_name="A3 mutant"
                    /note="created by site directed mutagenesis"
                    /citation=[3]
                    /phenotype="silent mutation"
  mutation          replace(13934..13936," ccg changed to ccc")
                    /gene="appA"
                    /standard_name=" P428 mutant"
                    /note="created by site directed mutagenesis"
                    /citation=[3]
                    /phenotype=" silent mutation "
  mutation          replace(13937..13939," gcg changed to gct")
                    /gene="appA"
                    /standard_name=" A429 mutant"
                    /note="created by site directed mutagenesis"

```

Figure 5 (continued):

```

/citation=[3]
/phenotype=" silent mutation "

DEFINITION pBluescript II XS(+) vector DNA.
ACCESSION X52327
VERSION X52327.1 GI:58061
KEYWORDS artificial sequence; cloning vector; expression vector; vector.
SOURCE synthetic construct.
ORGANISM synthetic construct
ARTIFICIAL artificial sequence.
REFERENCE 1 (bases 17732 to 20623)
AUTHORS Thomas, E.A.
TITLE Direct Submission
JOURNAL Submitted (20-FEB-1990) Thomas E.A., Stratagene Cloning
Systems, 11099 North Torney Pines Rd., La Jolla, CA 92037, USA
REFERENCE 2 (bases 17732 to 20623)
AUTHORS Short, J.M., Fernandez, J.M., Sorge, J.A. and Huse, W.D.
TITLE Lambda ZAP: a bacteriophage lambda expression vector with in
vivo excision properties
JOURNAL Nucleic Acids Res. 16 (15), 7583-7600 (1988)
MEDLINE 88319944
REFERENCE 3 (bases 17732 to 20623)
AUTHORS Alting-Mees, M.A. and Short, J.M.
TITLE pBluescript II: gene mapping vectors
JOURNAL Nucleic Acids Res. 17 (22), 9494 (1989)
MEDLINE 90067967
FEATURES Location/Qualifiers
Source 17732 to 20623
/organism="synthetic construct"
/db_xref="taxon:32630"
CDS complement (18967..19827)
/gene="Amp"
/product="b-lactamase"

BASE COUNT 5449 a 4847 c 4902 g 5424 t
ORIGIN
1 TCGAGAGTAT CTTTGTCAAG TGTGCTTCCA ACAAAAGGGGT ACTGTTGCC ACATAGAAAAG
61 ATCTAAACTA ATTAATTAAT CCCTCACCCG CAAATCTTC AGTCATAAG TTAGCACGAT
121 TGTGAACAA GTTCTCAAA GGAGAGATAAC AGATGAGTGC GTATAGGGTG GACCTGGCTG
181 CTGAGGGACAC ACCTGCATCT GACTAAGAAG AGCCACGGTG TTAGTTGAAT GGTGTGGAGT
241 AGGGTGGTTC TGTGGGACAG TAGAAAATCG AGAGGCATGT GCGCTTTAGT GAACTGATGG
301 AAGCTACCCC AAACGACAGA GATTGTCAGT CAGGCCAATC CGTTTCGAGT TTGATGGGCA
361 GCGGGACAGT GAGACAGACA CACCTACTCA GTTGGAGGAA GGATGAGAAC AATGCCAGC
421 AGGGATTGAG AGACCCCTGAC AGGCCAAGG CCCTAACACA CACACCTAAC ACCTCACTTG
481 ACAAAAGCTCC CAAAGACCAA AGACTTGTTC TCCATTAGAA ATGACAGCTG GCTTGACCOG
541 ACAGCATATAA AAGCAGACTG TACTCTGATT GGAGAACATT ATTGTGTTTC ATTCACTATT
601 ATAAAAGGAC AGTATTACAG ATTTTGTGT ACACTGCTGT TACATGTGGG GCAGTGTGTC
661 TTTAAGTAGG GTAAAGTACT CTTAAATAT GGGTCTTAA TATTTTTCC TTAACTCAA
721 GTCTCTTACT GTTTAAATGA TTTTTATTT GTTTAATATG GAGGAAAAAG AAGCGTAAAT
781 GGACAATATA TATTTAGAGA AAGATGGTTA GCTGTCAGAA AAATATGCAA ATCAAAATCA
841 CACCAAGACT GCAGCACACC CCTGTCAGAT GGCTGTGATC AAGAAAATAA ATGACAATGA
901 GTGGTGGTGA AGATGTACTA AAGGGAAACA CACACACACA CACACACACA CACACACACA
961 CACACTGGAG CAACCACTGT GGAAATCTGAT ATGAATGGTC CTCAAAACC TGAAGATAGA
1021 GCGGGGGCGTG GTGGCATACA CTTTATTC CAGCACTGGG GAGGCAGAGG CAGGTGGATC
1081 TCTGAGTTCC AGGCCAGCT GGTCTATAGC ACAGGTTCTA GGACAGCCAG GGCTACACAG
1141 AAAAACCTTG CCTGATTAAC ACCAAACCAA ACCAAACCAA ACCAAACCAA ACCAAACCAA
1201 ACCAAACCAA ACCAAACCAAG ACCAAACCAA AACACTGAAG ATAGAACTTC AGTATTCCAT
1261 TCCTAGATAT ATACCCAAATG GAGACTAAGT CAGCAAGACA CCTGCACAGC CATGTTCACT
1321 ACTACACTGT TCACCAACAGC CAGGCTGTGG AACCAAGCTG AGTGTCCATG ATAAATGAAT
1381 GGATAGGTAA CTTTCAAGGT AAATGGACTC TGCTGTGTAC ATGCCCTACA TTCTGTTTAT
1441 TCATTTTCTC TTATGAGGTG TCCATTCAAGG AGTCACATGG TAGTTCTATT TTCAGTCCTC
1501 TGAAGATACT ACACGTGTC CCACAGTTA CACTTTTATC AGCACTGAAT AAGGGTTCCT
1561 CTATCCTTAC CATCATTTGT TGTAAATTTT CTTGATGACC CTCCTTCTGA CAGGGATAGG
1621 ATGTAATATC AGTGTGAGGA AGTACAACCTT GTTTCTAAG TATTTATTTGG CCCCTTGAT
1681 TTCTTCTTTT GAAAACGTGTC GGTCTCTGAC ATCTGTCAG GTATTCATTG GATGTTGTTT

```

Figure 5 (continued):

1741 CTTTGGTGTG TGAGTTCTTA TGAATTCTAG ATGTTAAATC CCTGCCTGTG GTTCTCTCCC
 1801 ATTCTGTAGG CTGCCTCCTC ACCCTGGCAA TTGTTGTCCT TGTTTGCGAG AAACCTTTGA
 1861 CTTCATGGAA TCTCATTGTC CAGTTTCCC TCCTCTGCTA TAGCCTGAGC TAATGCACTG
 1921 GTTTTACAG AGCCCTGGTC TATGCTTAA TCCTCTGTG GCAGCTTCGG AGTTTCAATT
 1981 CTTACATTAA GATCTTGAT CCACCTTGAA CAAGTTTGG AGCAGGGTGA GAGATACGAA
 2041 TCTAGTTCA TTCTTCATA TGTGATCCTA GTTTACATAG CATCGTTGGT TGAAGAGGTT
 2101 TTATTTTATT TTTAATAAT GTGTCATAAA AAACGAGGTG GTTGTAGCAG TGTGGATTG
 2161 TTCTTTGTC CTTTGATCTA CAGGTCTGT TTGTTGTCAG TCTCATGATG TTTTATTGCT
 2221 ATGGCTCTGT CATACTAGT CAGGGTCAGGT ATTGTTGATAT ACCTTCAGTA TTGCTCCCTC
 2281 AGACTCAGGT TTGCTTGGC CAGGAGTCAT TTACTCTAGT GCTCTTAGAG CTCCCCCAGC
 2341 ATGTTAGCTGC TACTATTCTT AGTTGATAAA TCAGGAAACT GGGGCTCAGA GAGATTAACT
 2401 GTCTTAACT ACTTCTGGGG AGGTGAAACG TGGAGAACRT AAACCTGTGT TACCCCTGTAC
 2461 TGCTCCAGTA GCTGTCGGGT GCTGGCTAC AGCAAAGCAC CTATACTATA TATTACTCAG
 2521 GAGGTGGAAA AACTCAGCCT CCCTTGGGGT TCCCAAGCTC CCAGGGTGTCC AGTCACTGCT
 2581 GGAAACCTCA TGGAGTCTGA AAGGAAGGGT TGAGGGTACA TGGGGCAGCG ATGAGGAGCC
 2641 TGGGGCTGGG ATCTCCAAA CACCTGGATA TCCAGATGCC ACTGGGTCAAG GGGGAGTTGG
 2701 GAAACAGAGTT GGGATGTCCA TGACCTGTG ACAAGGCGAG GGCCAGGGGG AGGATAACTC
 2761 TGGCTTAACT AATTGCGAA AGTCTTCTAG TTAGCAGCAG TTGTCCTGGGA GCACAGAGGG
 2821 GCTTCTGTG AGAGGGTCAG GCAGTGCAGC TCTGTTAGCG AAGGTCTTCT CCATGTTCCC
 2881 CATGGTGGGTT CTTGATGAAA GAGACAGTCC TTGGTCTCAA ACTGGTTTAT TGATTGTTCA
 2941 TTGTCGAAA TGGGGTGCACA CCACCTCTC AGGGTGGACCC AGAGATCAA TACCTTTG
 3001 AGGGAGGAAT ATCTGGGAAG GGACGCTTAC TGGCTAAACC CTCAGGGCCT CTAGATACT
 3061 CATTAGCATG GAGAACTCTG TTCTGGGCTA CATGACCACCA GGCCACATTT CCACAAAGCCA
 3121 CATGTGGGAA GTGTGGCACA TGTCTAGGC CAGGAATCTG GTAGGGAGCG TGGAGCCACC
 3181 TACCATCCCA GGTGGGTGCG TGGGTGCCAG GGACCTGAA CCCGCTCAAC CTTACCAAGT
 3241 TTCTTGGCAG GGTCCACTGT CCTACACAGA AGCTGGAGGA GGTGTGAGGG TTGTTGCTTT
 3301 GTGGAATGTC CCATGTCGCT TGGGGCTCAG TTCTCCACCT TGTCACCTCAT TGGTTTGGGT
 3361 ATAAAAAAAGTG GGGATACTTT ATTATTCTCT GACTCGGTCC TGAGGAAAAAA GCATCGTGGC
 3421 AGTCCAGGAA CCACACCTCTG AGGTCTTCAG ACTGAAGGGG CTCCCTAAAGT CTCTGGAGTC
 3481 TCTCCCCCTC ACAGAGCTGC CAAAGTCTAG TTCTTGTGAG GGATAACAGA GCGATGCTTG
 3541 GTAAAGCAGAC AACAGCATTT GTTACTCAA CCTTCTTTG TCAGCTCCCT CTTCATAAAC
 3601 AAGTTGAGAC ACCATGCTGG CTTGAGGAAG ACTTCTAAAG CCAGACAAC GTGCAAGGAA
 3661 GAAGAAGAAG GGGCAAGTGG AGTTAGCCTG GATGTAGCCC TCAAAGTCTC CAGAGACCAG
 3721 CCATGAAGGC TCAAGTGGAG GGCAGAACCT GCAGCAGCCA AGCATCTGGC AGGAGAGGAT
 3781 CCTGGGAAACC CCTCTTACCAT GACACACATT CTTCCTGCAG GTCACACTTA ATAGGCCATT
 3841 TCTTATTGTC ATCTATCATG GTGTTCTGTG CGAGATTAAT GAGGTGTTAT GCTGCGAAC
 3901 GAAAGTTATA TAAAACAAG TCCCTTTCCC TTGTCACTGC TGCTAAGAAT GTAGCAGAAA
 3961 TTGTCCTCAAG TGTCTCTCTA ATCAGAAACA ATAAAGGTCT CTTGGATTCA AAGCCCTCCA
 4021 GTTCTCTCCT TCCTTGCTGA GCCTTGGACA CCCATACAAA CCTCTGGAT GCTACAGCTC
 4081 TGGGCAGAGA CTCCAAGGTG GGGAGAGACT GATGGTACAA AAGCAAATA CTTGTTTGGG
 4141 GGTACACCCCA CTOCTCTGCCC TGTGTTGTTCT CGCAGTCAG TCTGTCAGAC AGGCCCTCAG
 4201 TGGGTCTTCC ATGGGAACA CGCAGAGGGG GGCATGGAT GGGAAATACCC ACACCCCTGGT
 4261 TAGTTTACCC CGGGCATGCT CTCTGCTCTT CATCCCTCTT CTGCCCCCTG CCACGGCTTT
 4321 CTCTGCAGGA ATCATATCTT CATATTGGCC CACAGTGTGTT CTCCCTACCC TAGCTATGAT
 4381 GTTACTTTA GAGTGACCTT AGCAGGGCTG GTGGGAATGAG TTCTCTAGAC GCTCACGGAG
 4441 ATGCTAGGGG AGAACAGTC TCTAACTACT GAGGTACTAAG AGTTCTGGT GGTGTTCTCT
 4501 GCCTTCCCT TGTAAAGTC ACCCTGAAGT TAGTGCAGAA GAAATCAGAG CCCAGTCACA
 4561 GAGTAAATAT GGTCCTGAAG ATTCTTTTG AGTGCAGGAGA ATCCATGACA TTCAAGAGC
 4621 CCTCTTGTG CTTTAAGTC TTTGGGGTTG TATCTCTGC TTGATGATG TGTGTTGTT
 4681 TATCAAAGAG TGAGATGGTT ACATAAGAGG TGCTCTAAAG GACAGAGAGG ATTTCGAATT
 4741 GTGGCATGTG ACATCTCTAG GCCTTGTCTT GGTGCCAGGA GGAACAGTGTG CAGAAAAGAG
 4801 TAAGAGGTCA TTTCCTGGAG GCTGTCACTA TAGAGGAGAT CTTACAGTGC ATTCCCTCCT
 4861 CCAGGCCCCCTG CCTGAGGATA GACATGTCAGT GACTGCACCT GAAACAGAGG CTTGGGATGG
 4921 AGAGTTAGGT TCACAGAAGG GAGGGTGGGA GATGGATGCT TGCTGGGTTCT TGGGCTCTCAT
 4981 CACCACTCTG TGACCAAGGGC GTCAAGCCCCAT GTGCTTATTC CATAGCTTTC TTTGCTATG
 5041 TTCTACTGAGT GTGGTGTGTTG TTGGGACCCA GCAGAAGGCCA GTCCCCAGGCT GACAGCTGTG
 5101 GATACACAGG GCAGCATGAG GGTCTCTAGC CTGAAGCAGT CAGGCTGGCA GAAGAGAAAG
 5161 ACCAGCACAC ATTCCTTCAA CCAACTATGT CTGAAAAAC AAACATATTA TATCACATAT
 5221 ATTGCTTAA TGAGACAGCT AAAATGACT CGGGTAGCAT GACTCCAGT GGGGATATCT
 5281 GCAAGTGCCTA TGAGTGGCAG AGGGACAGCC AATGTGAGGC AAGAAGGAAT TCTGGCTCAA
 5341 CACAGCTTAG CTCCCTGGTG TTGGTTCAA CTTTGAGAGT TTGACCAACAA GCACCTTATT
 5401 TTGACATAT TAAACAGAG CACAACTTG GGAAAAGTT TTCTTATGAA ATTATCACA
 5461 ATAAAGCTTA AGGCATGACT ACATTTAAAT GCCTTGGCAA AGTATAATGTG CCCCTCTCCA
 5521 CAAGAATGGT TCTATTGACT GAGAAATAAT GTTCAGGATA AAGATCCAGG AAGAAAAGAT
 5581 CAGGGATAAG TAAATACTA AACTCTTTG CAAAGTACAT AGACCCCTTT TCATAACAAAT

Figure 5 (continued):

5641 GGGTTCTATT GACTGACAAG CACTGCTAG GAGTTGGAA AGAGTCTAGC ATAAGCACGA
 5701 TAGCCTGGAG ACTCTAGTGA GGCTAGTCT TACAGACAGC AAAATCACC AGGTTACAAA
 5761 CTACATTCAAT TCCAGTTTCTT CTGATCAGGC ACAGGTATGA ATCCCTCTG TTGAAGAGAA
 5821 AAGTCCATGT GTTTAAAATA TCTGGTTCTT CCAGTGTATAGGAGAAG ACTTGAGCCC
 5881 TATACAACCTT CCACCTGGAG TGACATCCTG TCTTCATGGT ATATTACATA CCTAGACACG
 5941 CTCATCTCAC AGACTTAGGA CTTTGTCTTC TGATCTCCAT TTCTGATCCC ATTCACCT
 6001 TTGCTTGTAGT AGTGTCTT TCTTCACTGC CTTGGTGACA ACCATGTAT CCTCTGTGTA
 6061 TTGAGTGTG ACCATTTCAT GATTTTACCT GTATGCAAGA TCACACAGTC TTGCTTTC
 6121 TGTCTGGATG CATGCTAATC TCTACACAAAC AACCCCTTCCC CGTCACTCG ATCTCCCTC
 6181 ATTAACACAT ACATGGTGCT GAAGAGGCTA GGGAGCTTC CTTAGTGGG GAGCTAGCTG
 6241 GCTATTGGC CTTTTGACT GTCCAGGAAG GCCCCCAATT GCTGAGACAA GAACCTAGAT
 6301 TCTTCATTAT TGACTCTAAC TCATGTATCA AGCAGAAGCT AATGAATAGT TATCAACAGG
 6361 ATCAGAGGT CCAGTGTAAAG ACACCTTGAC ATGAAAGAAC GGAGGAAGGA CAGATGGATG
 6421 CATAAAAGCA GGACCACTGC CCCAGGAAGG TCTTGGAAAC TGATGCGAGGG CAAAGGACAG
 6481 GTTATAAACAAATCCTTGGAG GAGTCAGGAA GAGCACAGAG GAGCTCAACC AACTGACCAC
 6541 TGCTTGGGG CTACCAACCC AATCTCCCT GTGGGAACAG CTAAGCTATC AGCCAAGGGT
 6601 ATAAACAGG CAGGACCTGT GGATGACATG GAGAGCATAG GGACCTGGG TCCAGCCCTT
 6661 AGCACCTGCA CTCTCAGGAT ACCTCCACAT TGTGCTTCTAG AGACCTTAGG GATACTGGGT
 6721 CCAGCCTTGT GTACCTTCAC TCTCAGGGTA CCCCCATCACT GTGCTTGGG GAGCCTAGGC
 6781 ACCCTGGGTC CAGCCTTCAG TACCTCGCCT CTCAGGACAC CCCACCATG TCTCTGGCC
 6841 CGTCTCTTCTC TCCCTCTCCT CCCTTCATT GTCTCTTCTC TGTTTCTTC TTGACTCTCC
 6901 TTTCCTCTCA CACCCCTACT CTAGTTCTCC CCTTCCCTCT CTGCTATCACC CTATTCTCTC
 6961 TGTGGTCCCT CCACCTTCTCT TTATCTCTCA TGCTTCTCTC CTCCCTCAAA TACTTGTCA
 7021 CCACATATACT TCAGGGGCCA GCTCTAGTGA CAAAGCTGT AATAGCAAGA CTCTCAGATC
 7081 TCCAACGGCT CAGAGGAGGCC AGACCCACCA AGAACCTCTC CCAGTCCAA TTTCAGGTT
 7141 CTTCGAAAGC TTTAGCAGAA TGTCAGGGG AGATGCCACT AACAGAAAGA TGCAAATTCC
 7201 AGTTGAGAGT GGGAAAGGCC CTTAGGTAGG TCCCATCTTC CAGGCCAAGG TCAGAGGGGC
 7261 TCTGTGTAAAT CCGGATGTAC AGGGCTCAGA ACAATGTTT GTTTTTAAGG TTATTTTATT
 7321 TTAGGTGTTA GTGCTTTCG TTGCTATGACC TTATGTGCAT CATGTGTGTG CAGGTTCTG
 7381 ATGACAGTAG AGGAGGGCTT TGAATCCCTG GGGATAGGAA GTTACAGGAA ATTATAAGCT
 7441 GCTTTGTGGG TCTTCCTAGT TTCCCAACAG AAGTGAATGC TCTTCACCCAC TGAGCCATCT
 7501 CTCTAGGCC AAGAGACATT GCTTTATGGA TATAATTGTG TGTGTGTGTC AACATTGAGG
 7561 AAAGGGAAAT AAAAAAAAATCCTCAGGCC TAAGGGTGTG CAGTTCTACT AATTGCTACT
 7621 TTAGTTGTG ATAAATGGC AGGTGCTTCA ACATTTATAT ATACAAAAC TTCCCTGCTG
 7681 GTGGTCAAC TGTGAGAACT GGGGTAAGTG GGTGAGTTCT CTTTTCTGT CTCTGCTCT
 7741 GTCTCTCTCC TTCCATTCTT TCTTAAAGGA AATAAACATT GCAGCTGGGT TATAGCTCAT
 7801 CAATATGGAA GTTACAGAAG TGAAAAAAGG CATTGCTTG GTGGGTGGG TTACCACTG
 7861 ATTTTGGGTT GTCCGTCAAG GAGGTCTGGG GACTGGCTGC TCTGCTCTG TCTGTATGAG
 7921 TGAGGGAAAGT CTGGGAGCA GATCCCTAA CCTTCAGGCCT GGCGCTGGTC CTGAGTGAAC
 7981 CCAGCCTCTC TGGTCTCTAGT AGCTTTTCTC AAACAGGAAT CTGAGTGTG ACAGGGAAACA
 8041 AGTACCCAGCC CATTGCTTAA GTGCCAGGGT TAGTGGGGG AGGAAGCTGC CATAGCTGGG
 8101 ATTAGTAGTT GTATGGATG TAGGAAGTCA TATCCTGGGA CAGCTAACTC TTAATGCTTC
 8161 ACTGGAGATT TTCAATGAGA AATTATCCC ACGGCCATA TGGCCCCATC CTTTGTCTC
 8221 CAACAGCCAA GTATTTTCCA TTAGAGGAGA CTTCTGTAC ACTTGATGGA TGCTCATTCC
 8281 AAGGTGACTT GGGGCAGTC GTACAGACTT GGGATGACCT CTGACAGCCT AACCTCTCCC
 8341 CAACAAAGGCC CCTCTATGTT TGCTATGTA TGTAATGTCA GACATTGTC GAGTGTCCG
 8401 CAGCACAGCC TGGCCAGTGT GAGGGCTCTC ATAGGTTCTC CACTGCTTA TCTACACAGG
 8461 GATAACGAGG AGGTAGCTG CAGTTCCAG TCTCACTTC CAGAGGAAGA GATAACCCCA
 8521 TCCCCAGGTCA TGAGGCCAGC AGTGGAAAGA AAGGAGGATT GAACCTAGGT CTTCCAAGTC
 8581 CCATTGATAG CATCTCTCA CAAGTCCCTT GCCACCCCTA CGATGCCCTA GACACTTGCC
 8641 TGCCCTTAT ACTAAGGAGA TGCAAGTACA AGGGTTTAC CCATGTAGCA GCTGAGGCAG
 8701 CTGGGATAG ATACCAAGCAG CAGGGCTGTAT GTCACTACTC TAACTCCAGC ATCCCCAGTC
 8761 TGTTCTCTG GAGTGTGAAA ATCCCTACTT AACAAGATG TGCAACAGTC CTTGGCTCTG
 8821 TGACCCATAG CTGGAAACAG GATTCTCATT GATTGTGGA ACATGGTGGC AGCCAGCCAA
 8881 AAAGAGGGTC TGCTACACAGA AGACACGTGT GGCAAGGCC CAGCAGACTC TGACTACCTT
 8941 AGCTTACAGA ATTACAAGGT CATAATGTCC TCTGCTTGG TCACCTCATG TTAAGGACAG
 9001 GCCCTAATGA AGATGGGCCA GAAGACTGAA GGAATGGCCA ACCAATAACT GCCCCAACTT
 9061 GAGCCCCATC CTACAGGCAA GCATCAATTCT CTGACACTAC TAATGATACT CTGTTATGCT
 9121 TGCAGACAGA AGCTCTAGCAT AACTATCCTC CGAGAGGTCC ACCCAGCAAC TGACTGAAAC
 9181 AGAAAAAGAT ATCCACAGGC AACAGTGGA TGGAGGTCAAG GAGTATTAT GGGAGAGCTG
 9241 TGGGAGGT TAAAAACCTT GAAGGGCTA GGAACCCCAAGGAGACCA ACAGAGTCAA
 9301 CTAAGAGACC TGTGGGAGCT CTCAGAGACT GAGCCACCAA CAAAGAGCA TACACAGGCC
 9361 GGTCCGAGGC ACCTGGCAGC TGTGAGGAG ACATGCAGCT CAGTCTCCAT GTAGGTCTC
 9421 CAATAACGGG TAGCCTGACT GCAGTATCCA ATCCCCAACCA GGGCTGCATA GTCTGGCTC
 9481 AGTGGGGAG GATGCCCTA ATCCCTGCAGA GACTTGTGTA GTGGAGAGCT ATCCAGGGGG

Figure 5 (continued):

9541 AACCCACCCCT CTCTGAGAAG GGAATGGGG A TGGGGGAGGG ACTCTGTGAA GAGGGGACAA
 9601 GGACAAACAA GAACCTAA TAGGTCAAGC CCTAAAGGCT TGCTAAGTAG CAGTGGCCA
 9661 GCTCTGTCT GTTCCTCAGC CCAAGGCTCA GCTCCCACCT GTTCTGTGT TTTCTGGCT
 9721 TTTCATGGC CTAGGACTTG GTGACCAGTT CAAACAATGG GGCCTGTGGA AGACACAATA
 9781 TACAAGACTA GGGACATTCC TGTTCGTG ACTATCCATA GCCTGATGTA GGTGGAAGGA
 9841 CCCAATCACT GGATTCTAC CCTTGACAA CCTTGACAGC TGAGGGCTC TCAGAAACCT
 9901 ATTTCTTCA CTGAAAATG AGACTCTCAA ATGAACGTC TGACAATCAT CAGGCTTATT
 9961 AAAGAGGTGT ATCTAACCTG AATGGCAAGC AGACAGCAGG CAAATGTCTG TATCAACCTC
 10021 TAGGAAGGAC AAGAACTGCT CACTGCTGCC CCCCAGGAGG CCATTGCTG AAACAGCTGC
 10081 TCTCTGTG TGTCACAGG CCTGCCCTCT CATTGCAAGC ACAGCCCTT CCTGTCTGAA
 10141 CCTCTGTCA GGTCACTGGG AAACAGATCA AGATGGAACA GGACAGCTCC TGATGGTAA
 10201 TAAAAAACAG TGGTCATGGC TATTCTAGG GTTCTATGCT TCTTCAGTCC ACACITGTGAA
 10261 GAGCTGTGGG CATGAACAC AGTGTTCGAG GTAGAGTGG GGTCTGAAA TTACACAGTGG
 10321 GGTGAGCTCA GTAAATGTGA GCTGGAGGTC ACTCGTGAAGA CACACAGTCC TGCTGCTTCT
 10381 GTTCCAAATA TCCTGAGGAG ACCACACATC TACTTTGTC ACAGGCCACA GTCTAGTTGA
 10441 CCTGAGAGTT ACCAGTTCT TATTGTGTG TGTGTGTGTG TGTGTGTGTG TGTGTGTGTG
 10501 TGTTGTCTGT GTGTGAGTGC AGGTGCACAT ATGATAGCGT ACACGTGAG GTCAAGGAG
 10561 AACTATCAGG CGTTGTCCCC TCCTACTTT CCTOGGACTC TGGAGAACAA ACATGGTCC
 10621 TTATTCCAGG GGAGCAAGTC GCTGTGAGCT GACACATCT GCTCACATAC ATTTACCTA
 10681 GACAATGGAG CCTCCATCAG AGTATTACTT TAGCTCTCA CCGATGGCAA TGCAACACCT
 10741 CTCTACCCAC ATAGGAGTTG GGTCTCACCA CACCCCCACAC CCCCCCTTCAC CAAACGTTT
 10801 TCAGTTACTT TATCTGGTAA AGTCTCATCAG AGAATGAAGC CAGTATTAAAG AACATGGAAT
 10861 CATTGGGAA CCTGGATCTA GCAATACCCCC ACCCTAGATG GAGTTGCTGA GTTTTCACCT
 10921 CAGATTATAA TTCCCCCTTA GCTTCTATGG TTATTCCTGA AACCAGGGGA ACTCGATTCC
 10981 TCCCTTGGGAA CCACAGACAT CCTGGCTTGT GAATTACAT GTCTACT TCTAAATCCAT
 11041 TGGTAGTAAG TGGTCACAG AGACACACTA CAGTCATGGC CAATGTCAGA GTAGGACAGA
 11101 TGTGAATCAT TCCCCCAGTC CTGCTGTTT CATGACTAAC CCTOCTCAGC ACAGTGACCA
 11161 TGAACCTACT TTCCCCCTTC TTATTTTTT AGAATGCTG GAATTTCCTA TTTTGAGAAA
 11221 TAATGCCTT GGGCAGCATT AAACAAAATC ATCTAGAAAG CTGGTTAAA ATACAGATGG
 11281 TTGAGTCAGT GAAAGAGTGA GGAATGTCAG TATTGGCCCCC TCACAGAGGC TGGCTCACTC
 11341 CAGCAGAGGT GGTTGAAGCT CTGGACACG GGTCAGGTGC ATAGGAAAGG TNGTCTGGGA
 11401 CACTGAGAAC CACAATTGAA CAAACAGAAC TGTTGGCTT TTTTTTTA AATGAGTTCT
 11461 CAAAAAAATGA CTGGCTAGCT TAGGCAAATA CTTOGAGCCA ACCCAACAGA ACATTTTCC
 11521 ATTGATTCTAT TCTGGATCTT CTTCTAGAC AATACTGAAC TGACCCCTG TTGGCAGTCT
 11581 CAAGTTTGAC AACATAGGGC TTGAACTTG GCACAAAGTC CATCACTGTC ACCCAAGCAT
 11641 CCTGGGTGAC TTGTTGGTTC GAATATCTTG GCTAACCTTA GATATTTCT TTGGAGTATC
 11701 TTAGAACATC CCAGGAAATA GGGCTTGTATT CTCTACCTGG GACCACAAATA TAAGTCACCC
 11761 TAGAATCCCA GGAGATCGTG CAGAGAAACA AGGATCTCTC TCGTGTGCT CTTCTTCAA
 11821 AGCAGTGAAGT AGTCACTCCA CTAACACTGAG TTCCCATCTG AGAGTCACCA GGAGGCTTGG
 11881 GGGCAAGGAAG CAGAGGGAAAG GCACGTGTTG TGTTGGTAA GTTTTGACTC TAACAAATT
 11941 GAAGACATAG ATGACATTGT GTCAGACTAA CAACAACCTA GACTCATGIG GTTTCTGTTT
 12001 AGGGATCAGA TTTTATTCTAT CTTGACTTG TCTTAGTGTG TAGAGAAAGG CTTCTACTG
 12061 GAGTGTAGGC TCAATAATGA CAGAAGAGAT AGCTATTTCC CCTAGGGACT GTGCTGCTCC
 12121 AAGTTGGTG GAGAAAGGC GTGGGGAACC TAGATGTGCT CTCTGGGAG GGGGCTGAA
 12181 GCTGGCTCA TAGAAGGTGT GAATTTTGC TGAAACATCT AAACAGAATT ATAGCTTATTG
 12241 AAAGTGAGCA GGCAGGGCAG GGAATGTGTT GCTATGTAT ATGTCATGAA ATATATTATG
 12301 TTATAGATAC ACACACATT GAACTCATT TGCAAGATGAC AGAAAATAGG TTATTTGCC
 12361 TCTCTTAAC TCTAAGCACA ATGACTTCCA GTTOCATOCA TTCTCTGAAA TGCCACAAATT
 12421 TCATTTTCA TTGTTGGCTGA ATAAAATTC ATTGCAACT GGGCCCTACT TCATCCACTC
 12481 CTGAGGGCAG GCATATCCCC TGGCTCCATT TCTTACCTAT TGTAAGAGAGA AGTGCACACTG
 12541 TCTTGTGAA AGGCAGCGT GAGAGAGGCA GGCACAAATT GTGGTTTT GTTTCTTCTT
 12601 CCTGCTATGA CTCTCCATT GTCAAGACCA AAGATCGATA AAAGCCGCCA CCATGAAAGC
 12661 CATCTTAATC CCATTTTAT CTCTCTGAT TCGTTEAAC COGCAACTG CATTOGCTCA
 12721 GAGTGAAGC GAGCTGAAGC TGGAAAAGTGT GTGATTGTC AGTCGTCATG GTGTGCGTGC
 12781 TCCAACCAAG GCCACGCAAC TGATGCAAGA TGTCACCCCA GACCCATGGC CAAACCTGGCC
 12841 GTTAAAATCT GGTGGCTGA CACCGCGCGG TGGTGAGCTA ATOGCCTATC TCGGACATTA
 12901 CCAACGCCAG CGTCTGGTAG CGAGCGATT GCTGGCGAAA AAGGGCTGCC CGCACTCTGG
 12961 TCAGGTCGGG ATTATGGCTG ATGTCAGCGA CGTACCCCGT AAAACAGGCG AAGCCTTCGC
 13021 CGCGGGCTG GCACCTGACT GTGCAATAAC CGTACATACC CAGGAGATA CGTCCAGTCC
 13081 CGATCCGTTA TTTAATCTC TAAAATCTG CGTTTGCCTAA CTGATAACG CGAACGTGAC
 13141 TGACGOGATC CTCAGCAGGG CAGGAGGGTC AATTGCTGAC TTTACGGGG ATCGCAACAC
 13201 GGCGTTTCGC GAACTGAAAC GGGTGTTAA TTTCCGCAA TCAAACCTGT GCCTTAAACG
 13261 TGAGAAACAG GACGAAAGCT GTTCATTAC GCAGGCATTA CCATGGAAC TCAAGGTGAG
 13321 CGCGACAAAT GTCTCATTAA CGGGTGGCGGT AAGCCTCGCA TCAATGCTGA CGGAGATATT
 13381 TCTCTGCAA CAAGCACAGG GAATGCCGGA GCGGGGGTGG GGAAGGATCA CGGATTCA

Figure 5 (continued):

13441 CCAGTGGAAAC ACCTTGCTAA GTTGCATAA CGCGCAATT TATTGCTAC AACGCACGCC
 13501 AGAGGTTGCC CGCAGCCGCG CCACCCCGTT ATTAGATTG ATCAAGACAG CGTTGACGCC
 13561 CCATCCACCG CAAAACAGG CGTATGGTGT GACATTACCC ACTTCAGTGC TGTTCATCGC
 13621 CGGACACGGT ACTAATCTGG CAAATCTCGG CGGCGCACTG GAGCTCAACT GGACGCTTCC
 13681 CGGTCAAGCCG GATAACACCG CGCCAGGTGG TGAACGGGTG TTTGAACGCT GGCGTGGCT
 13741 AAGCGATAAC AGCCAGTGGA TTCAGGTTTC GCTGGCTTC CAGACTTAC AGCAGATGCG
 13801 TGATAAAACG CCGCTGTAT TAAATACGCC GCGCGGAGAG GTGAAACTGA CCCTGGCAGG
 13861 ATGTGAAGAG CGAAATGCC AGGGCATGTG TTGTTGGCA GTTTTACGC AAATCGTGA
 13921 TGAAGCACGC ATACCCGTT GCAGTTGTG AGCTTACCCGG GGATCACAAAC TTGCCCCCTCG
 13981 AAGAGGAAC ACAGAAGGAT CGCACAATC TCCGTCTGC TACTCTCCAG TGGTTTCATC
 14041 TTACTCTGA TGGCATTTCC CTCTAGAAAG TGCTACTATC ATCCACACAT TTCTACCTGA
 14101 GACCACCCAA AGGACCCCTCC CAAATTCTCT TCCCTCTGA GTAGTCTCCA CACTGTTAC
 14161 CACCATCCCA GAATTTAAAT CCTAACTGCA CTCGGCGTG TGACTTGCCT CAGTCCCTGC
 14221 AATAAGAGTT GTGGCAGTG CCAGGCGTGG TGGCGCACGC CTTAAATTC AGCACATTGGG
 14281 AGGCAGAGGC AGGCGGATT CTGAGTTCGA GGCCAGCTG GTCTACAGAG TGAGTTCCAG
 14341 GACAGCCAGG CCTATACAGA GAAACCCCTGT GTCGAAAAC CAAAAAAA AAAAAAAGTT
 14401 GTGGCAGAG TGTGGTTAT ATACCAGGTG GAGATTCAA ATGAGTGGCT GAAGCTGTAG
 14461 CCAGAAGGAA CTTAGAGGAT AGCTCATAAC TTTAAAGAA ATGTAGAGAG TAGCAGAAAC
 14521 ATTGAGAGGAC TGGGCACACA CGCACTGTG GAATGTGGCA GAACACAATC CAGCCAGCTA
 14581 TACATGCATA AGTGTATATT GGCGCCATCC TGACTGTGA GACACAGGAA AACAGATAGA
 14641 CGGGGTTAGG TGGCCATGGC CTTTCTGCC TGCCCTCTCC TAAGGGTCACT CTCAGACCT
 14701 TATGCTCTCT TAACCTCTCC ATTGCTACTT AGCTTCTAGA TATCACCTCC AGATTAGTCT
 14761 CCTGGGTAC ATCAGTGATC CTGGTGTATAT CCAGGGCTTC CTGATCCAT CTTTGTCTA
 14821 GAGGCTGCAA CTAAAGAGGT CTCTTAAATA CTTCACACCC TGATGCAAA AGGAAGACAC
 14881 AGAAGTTCAC AGAGGTGAAG TGATTCACTG AGGACATACA GTGAGCAAGC ATCAGGTCC
 14941 GGATTATCTG ACTCTACTCT AACTTTTATG TAAATGTGCT TTATGCCATT AACACTGTCA
 15001 TTCTGTGCT TCAGCTCTGG GAGACTCCCCA AGCACTCTTA GGCACAAGCC ACAATTAAAGG
 15061 GACTCTGACA CTCTGCATTG ATTAAATTAGC ATGGTGGTCT CTATGTTCC AGATTCACTGA
 15121 TTGTTTCACT TTCCATATAG GCTATGAAGG GTGTGAGGAA ATTTTTGGG GACAGAATTG
 15181 GAGGCAATCC ACCTCTCTCA GGAAGCTCT ATCTGGAAA CCTTACAAC CAGGGACAGT
 15241 AACTGTAGGC CCAGTCCTTG GTGTCCAAA TGGGTTTATG GTTTGAATC TGCAAAGCCT
 15301 TCCATGTGCT CAAAGGTTTG AACATGGAGC CTCCCTCTGG TAACACTGTA TTGGAGGCCT
 15361 TTGAGACTGG ATGCTCTTGTG GTCCCATGTT TGCTACATC ATCTGTCAAG ATATGACCC
 15421 GGCATGCTAC CAGCTACCCAC AGACTATGCC TCTCCAGCTT TCATGTTCTC CCCACCATGA
 15481 TAGACTTGTAA TCTCTTAAATGGAATCAA AGCAAACCTT TCCITGCATTA AGTTTTTTT
 15541 TTCTGTAAATG GTGTTGGTC ACAGGGACAA GAAAACACTC AATACAGATA ATTAGTACCA
 15601 GAGTTGAGGT TCATGTTCTG AGCAAGTTGG ATCAAATTG TAGGGTTTG GAACIGATT
 15661 ATAAGAGACAA TGTAGAAGAG TCTGAAGCTG TGGGCTACAG AAGTGTCACT AGTTTTAAG
 15721 AATAGTTTAA TACACCAGG GAATTGTGAA AATCAGATG CTCACACAAA GGCAGACAGG
 15781 AAAACGTGAG CATGTGGCGT GTGAGAGGGC ATAAGAAGGA ACCTAGGGGG AAATGAGCTA
 15841 GAAGCCATTC GGCTACGTTA GGGAACGTGT GTGGCTGTG TTGGCCCATG CCCCCTGGCAAT
 15901 CTGAATGAGG CCAAATTTTA AAGGAGTGGA CTAACCTGAT TGTCAGAGAA AATATCAAGA
 15961 CAGCACCCA CTCAGGCTAT GCGCTGTTTG TGACCGACCA GCTACTCTTA GCCAGCTCTA
 16021 TTGTGAAAT CCAGAGCAAT TATCAGAGCA TGAAGATACA TACAGTTAG TGAAGTAAGG
 16081 GGTGTTGGTC CCTAAGTGGA TGGTCATAA ATCTATGTAG GTGATGCCAT AGTGCACACTT
 16141 GATAATCCAA AATATCAGCA AATGTTGGATG TCTTCCAAGG AGACCTGTAG ACACACATT
 16201 TAGAACTTTG CTCATGGCTG TAAATTTAAG CTAGCTAGAA ATCAATTCT GAAGAGGTTA
 16261 GTCTGAGTAA CGGTTCCAGG GCAAACATTC AGTGTGGCA AGGAAGGCAAT TGCAGTCAGG
 16321 AGCCAAAGGT CAGCTGGTCA CATTGCATCA AGAGTAGAGA GTCAAGACTGT GAGTAGAAAG
 16381 AGGATACAGG TTATAAAACC TCACGTGCA CTCTCAGCAA TCCATTCTCT CCTAAAAGGC
 16441 TTACCTCTTAAAGATTTA GTCTTCAAA CGAGTACCG TAGCTGGGA ACAAAAGTTG
 16501 AAACAAATGA GCCTTGTGG GGCATTTCACT TCTTAAACCA GGGCATCACCC TAGGAGGAGC
 16561 CCTGTGTGCA GTAGGAGTGT TGGCTCTGT GTCAAGGAATG CTCAAGCTAA TAAGGGTCC
 16621 TCTATCTGAG GGACCCATATG AAGATTCAAC AAGTAGTTGT GAGAATTCC TGTAAATGGA
 16681 TGCTACCAAT TTGACATTG TAGACCTGCT ATTGTGTGCT TCTTATTTGG GCTCTCCCAT
 16741 CTCCCAACTT TCCAAACCCAT ATTCCACATT AATCCCCTCC ACCACCATGC AACACTAGGT
 16801 AGGAGAGAAG GAAGGTAGA AGAGAAAAGTG GGTATAGATC TATTAGACT ACTTCCCTGCT
 16861 GATTAGGGC AAGTCCAATC GTCAATTGCA GGATACCTCC AACAGCAAC CAGCAAACCA
 16921 GCAAATCAGA AACAGCRAAA GCAGCCAACA AGGCAGCACT AACAGCAGG ATTGGGGTCG
 16981 GTAGCGTGGG AGCAGTCACT ACTGGCTTC TCATGGCTTT GGCATTAATA CTCTCTCAAG
 17041 AAATTCCGTA ATTTTTCCC CACCACTGAA AATCCGTAAT TTTAAATGC AACTATCTA
 17101 CAGCTGGCAA AAATCAGATC TCTCTTAGAG CACAAGACAA ATCATAGTAA CTGGCTATT
 17161 GCAATCTGAA GCATCTCAAT ATCCCACACC TGGGATTAAA ACAAAACAT ATTCAACATCA
 17221 CATAACTGTT TTTTTTCTC AATTTTTTACT TAGGTATTTC CTTTATTAC ATTTCAAATG
 17281 CTATCCGAA AGTCCCCAT ACCCTCCAC CTCCCTGCTC CCCTACACAC CCACTCCAC

Figure 5 (continued):

17341 TTTTTGACCC TGGAGTTCCC CGGTACTGGG GCATATAAAG TTTGCAAGAC CAAGGGGCCT
 17401 CTCTTCCCAG TGATGGCGA CTAAGCCATC TTCTGCTACA TATGCAAGATA GAGACACGAG
 17461 CTCTGGGGGT ACTAGTTAGT TCATATTGTT GTTCCACCTA TAGGGTCGCA GACCCCTCA
 17521 GCTCTTGGG TACTTTGCT AGCTCCTCCA CTGGGGGCTC TGTGTTTAT CTAATAGATG
 17581 ACTGTGAGCA TCCACTTCG TATTGACAG GCACTGGGCT AGCGTCACAT GAGCCAGCTA
 17641 TATCAGGGTC CTTTCAGCAA AACCTTGCTG GCATGTGCAA TAGTGTCTG GTTTGGGTGGT
 17701 TGATTATGGG ATGGATCCAC TAGTTCTAGA GGGCCGCGCA CGCGGTGGA GCTCCAGCTT
 17761 TTGTTCCCT TAGTGAGGGT TAATTGCGCG CTGGCGTAA TCATGGTCAT AGCTGTTCC
 17821 TGTGTGAAAT TGTATCCGC TCACAATTCC ACACAACATA CGAGCCGAA GCATAAAGTG
 17881 TAAAGCCTGG GGTGCTTAAT GAGTGAGCTA ACTCACATTA ATTGCGTGC GCTCACTGCC
 17941 CGCTTTCCAG TCGGAAACC TGTGCGCCA GCTGCATTAA TGAATCGGCC AACGCGCGGG
 18001 GAGAGGGCTT TTGCGTATTG GGCGCTCTTC CGCTCCCTCG CTCACIGACT CGCTGGCCTC
 18061 GGTGTTCCGG CTGCGGGAG CGGTATCAGC TCACCTCAAAG GCGGTAATAC GGTTATCCAC
 18121 AGAACCGGG GATAACCGAG GAAAGAACAT GTGAGCAAAA GGCCAGCAAA AGGCCAGGAA
 18181 CGTAAAGAG GCGCGTTGC TGGCGTTTCC CACAGGCTC CGCCCCCTG ACGAGCATCA
 18241 CAAAAATCGA CGCTCAAGTC AGAGGTGGCG AACACCGACAG GGAATATAAA GATACCAGGC
 18301 GTTTCCCTG GGAAGCTCCC TCGTGCCTC TCTGTTCCG ACCCTGCGC TTACCGGATA
 18361 CCTGTCCGCC TTTCTCCCT CGGGAAAGCGT GGCGCTTTCT CATAGCTCAC GCTGTAGGTA
 18421 TCTCAGTTCG GTGTAGGTGCG TTGCGCTCAA GCTGGGCTGT GTGACGAAAC CCCCCGTTCA
 18481 GCGCGACCGC TGCGCTTAT CGGGTAACCA TCGTCTTGAG TCCAACCCGG TAAGACACGA
 18541 CTTATCGCC CGCGCAGCAG CCACGTGAA CAGGATTAGC AGAGCGAGGT ATGTAGGGGG
 18601 TGCTACAGAG TTCTTGAAAGT GGTGCGCTAA CTACGGCTAC ACTAGAAGGA CAGTATTTGG
 18661 TATCTGCGCT CGTGTGAAGC CAGTACCTT CGGAAAGAAA GTGGTAGCT CTTGATCCGG
 18721 CAAACAAACCGC ACGCGTGGTA GGGGGTTT TTTGTTTGC AAGCAGCAGA TTACCGCGCAG
 18781 AAAAAAAGGA TCTCAAGAAG ATCCTTGTAT CTTTCTACG GGGTCTGACG CTCAGTGGAA
 18841 CGAAAACCTCA CGTTAAGGGG TTTGGTCAT GAGATTATCA AAAAGGATCT TCACCTAGAT
 18901 CCTTTAAAT TAAAAATGAA GTTTAAATC AATCTAAAGT ATATATGAGT AAACCTGGTC
 18961 TGACAGTTAC CAATGCTAA TCAGTGAGG ACCTATCTCA GCGATCTGTC TATTTGTTTC
 19021 ATCCATAGTT GCCTGACTCC CCGTCGTGTA GATAACTACG ATACGGGAGG GCTTACCATC
 19081 TGGCCCCAGT GCTGCAATGA TACCGCGAGA CCCACGCTCA CGCGCTCCAG ATTTATCAGC
 19141 AATAAACCGG CGAGCCGAGA GGGCGAGCG CAGAAGTGGT CCTGCAACTT TATCCGCCCTC
 19201 CATCCAGTCT ATTATTGTT GGGGGAAGG TAGACTAAGT AGTTGCGCAG TTAATAGTTT
 19261 GCGCAACGTT GTGCCATTG CTACAGGCAT CGTGGTGTCA CGCTCGTGT TTGGTATGGC
 19321 TTCATTCAAGC TCGGGTTCCC AACGATCAAG CGAGGTACA TGATCCCCCA TGTGTGCA
 19381 AAAAGCGGTT AGCTCCCTCG GTCCCTCGAT CGTTGTCAGA AGTAAGTTGG CCGCACTGTT
 19441 ATCACTCATG GTTATGGCAG CACTGCATAA TTCTCTTACT GTCATGCCAT CGTAAGATG
 19501 CTTTCTGTC ACTGGTGAGT ACTCAACCAA GTCAATTGTA GAATAGTGT TGCGGCGACC
 19561 GAGTTGCTCT TGCCCGGGCT CAATACCGGA TAATACCGCG CCACATAGCA GAACTTTAAA
 19621 ATGCTCATC ATTGGAAAAC GTCTCTGGG CGGAAAACCTC TCAAGGATCT TACCGCTGTT
 19681 GAGATCCAGT TCGATGTAAC CCACCTGTG ACCCAACTGA TCTTCAGCAT CTTTACITTT
 19741 CACCAAGCGTT TCTGGGTGAG CAAAACAGG AAGGCAAAAT GCGCAAAAAA AGGGATAAAG
 19801 GGCGACACGG AAATGTTGAA TACTCATACT TTCTCTTTT CAATATTATT GAAGCATTAA
 19861 TCAGGGTTAT TGTCTCATGA CGGGATACAT ATTGATGT ATTTAGAAA ATAAACAAAT
 19921 AGGGGTTCCG CGCACATTTC CGCGAAAAGT GCCACCTAAA TTGTAAGCGT TAATATTTG
 19981 TTAAAATTCG CGTAAATTT TTGTTAAATC AGTCATTTT TTAACCAATA GGCGAAATC
 20041 GCGAAAATCC TTATATAATC AAAAGAATAG ACGAGAGATAG GGTGAGTGT TGTCCAGTT
 20101 TGGAAACAAGA GTCCACTATT AAAGAACGTG GACTCCAACG TCAAAGGGCG AAAACCGTC
 20161 TATCAGGGCG ATGGGGCACT ACGTAACCA TCAACCTAAT CAAGTTTTT GGGGTGGAGG
 20221 TGCGTAAAG CACTAAATCG GAACCTAAA GGGAGCCCCC GATTTAGAGC TTGACCGGGGA
 20281 AACCGGGCGA ACGTGGCGAG AAAGGAAGGG AAGAAAGCGA AAGGAGCGGG CGCTAGGGCG
 20341 CTGGCAAGTG TAGCGGTCACT GCTGCCCGTA ACCACACAC CGCCGCGCT TAATGCGCCG
 20401 CTACAGGGCG CGTCCCCATTG CGCATTCAGG CTGCGCAACT GTTGGGAAGG CGATCGTGTG
 20461 CGGGCCTCTT CGCTATTACG CCAGCTGGCG AAAGGGGGAT GTGCTGCAAG CGGATTAAGT
 20521 TGGGTAACGC CAGGGTTTC CCAGTCACGA CGTTGAAAAA CGACGGCCAG TGAGCCGCG
 20581 TAATACGACT CACTATAGGG CGAATGGGT ACCGGGCCCC CCC

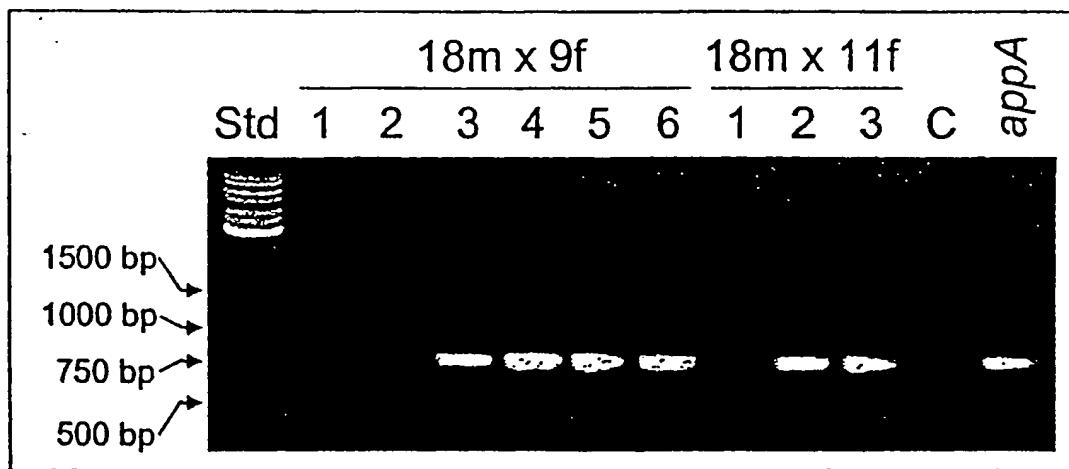


Figure 6

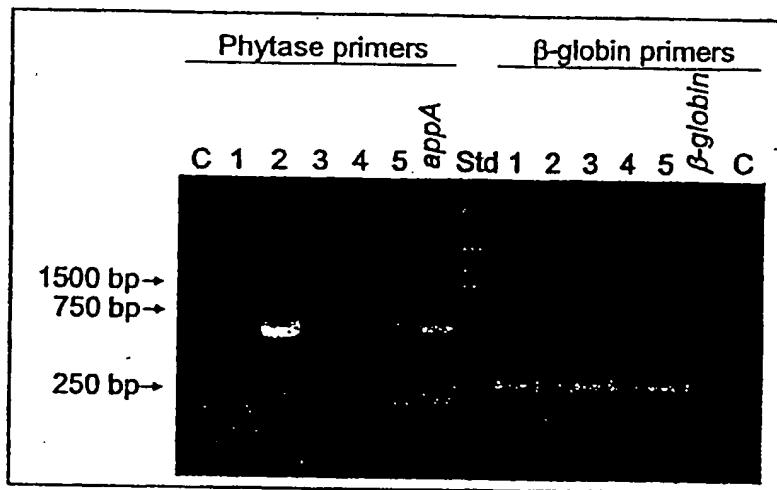


Figure 7

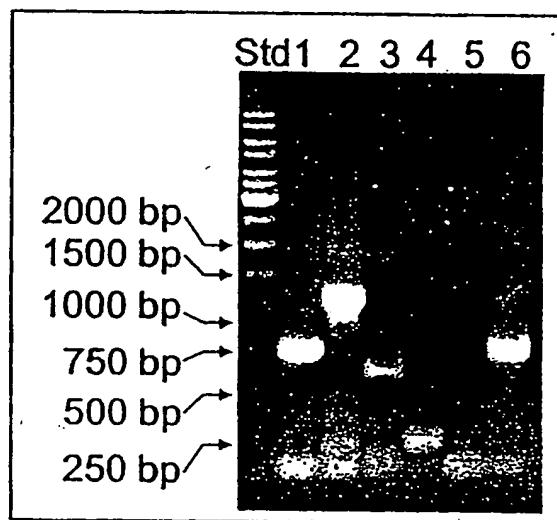


Figure 8

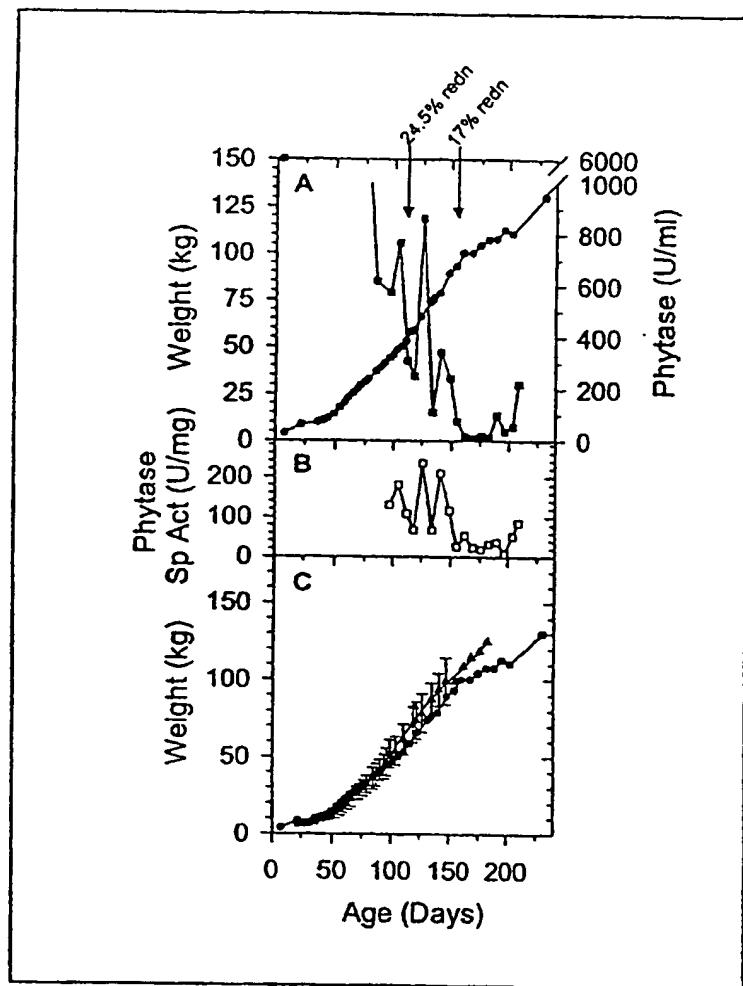


Figure 9

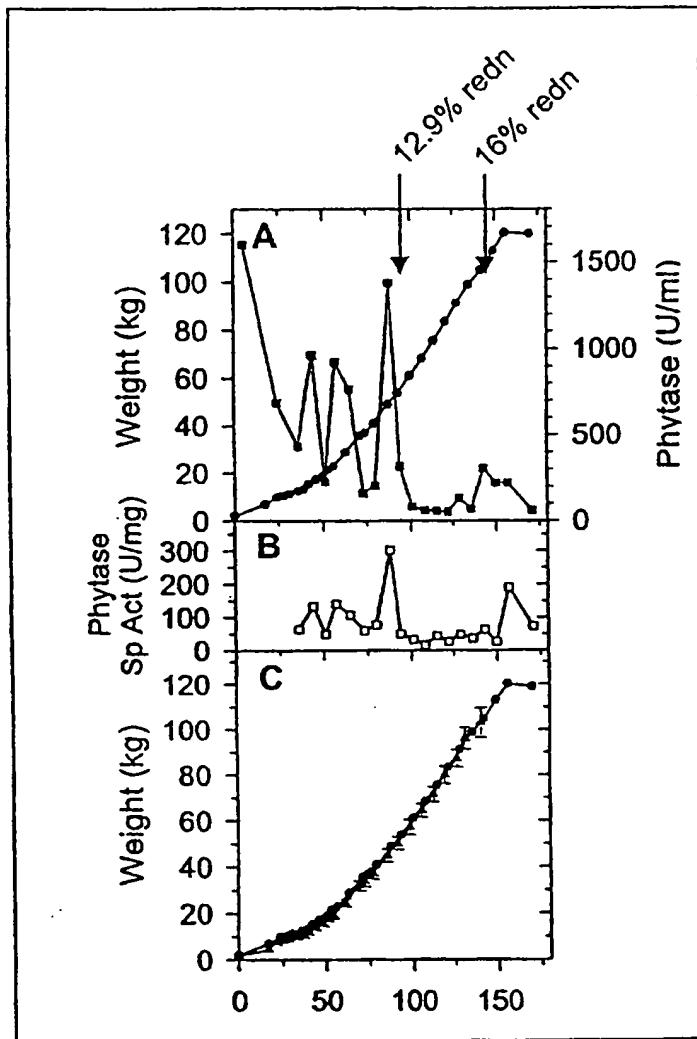


Figure 10

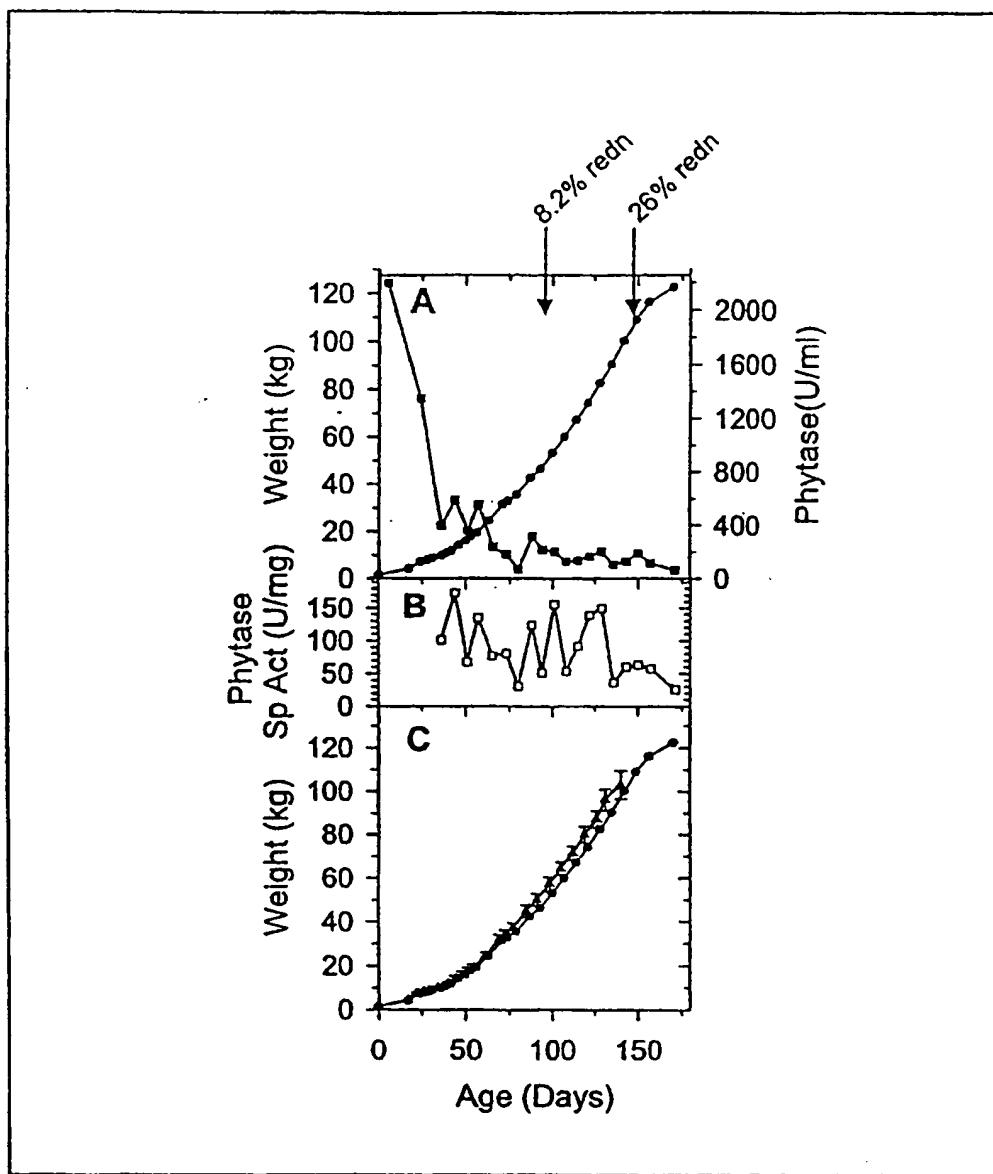


Figure 11

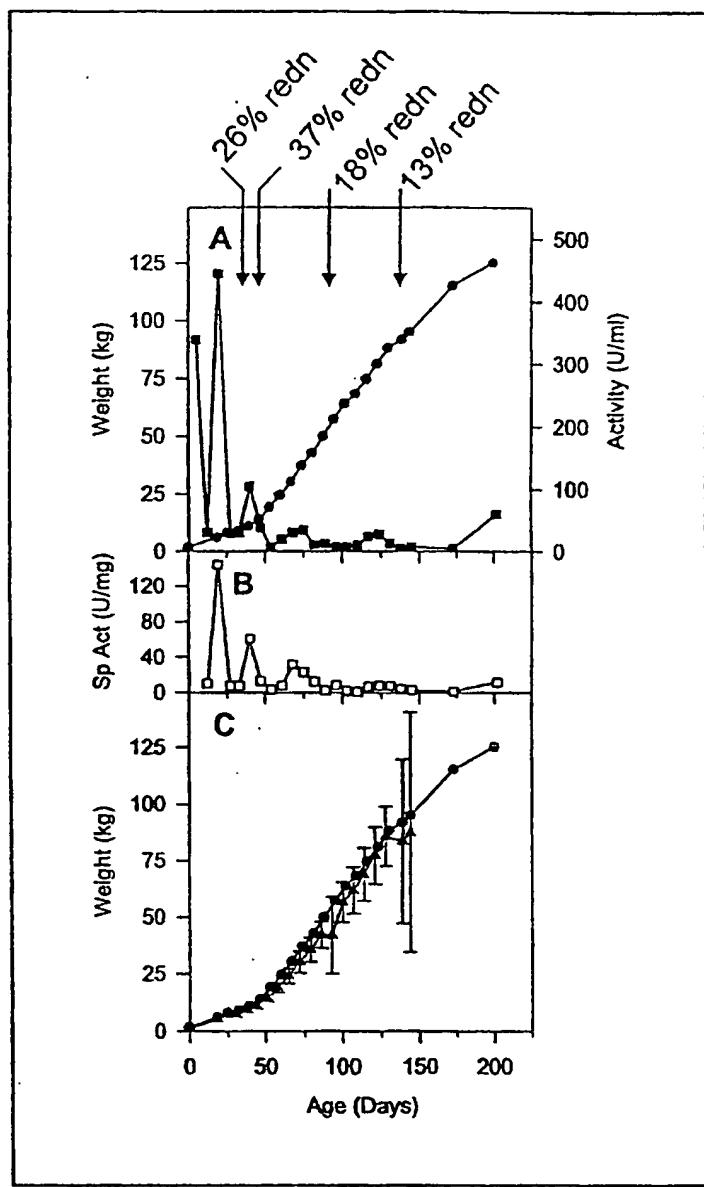


Figure 12

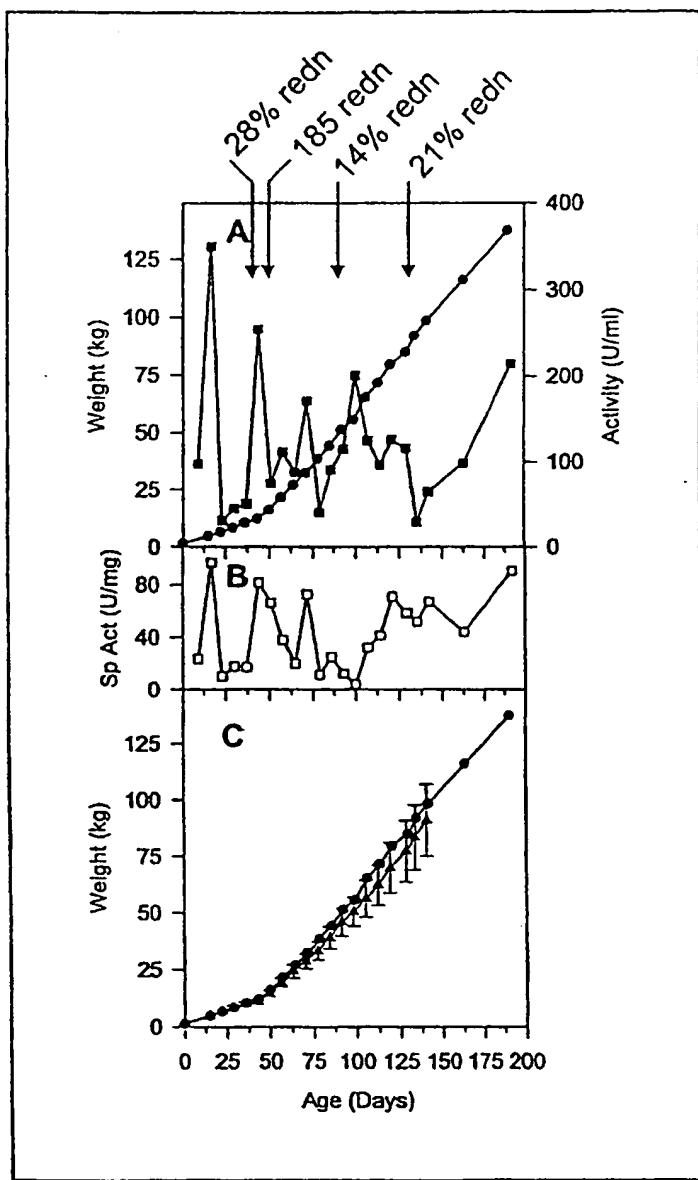


Figure 13

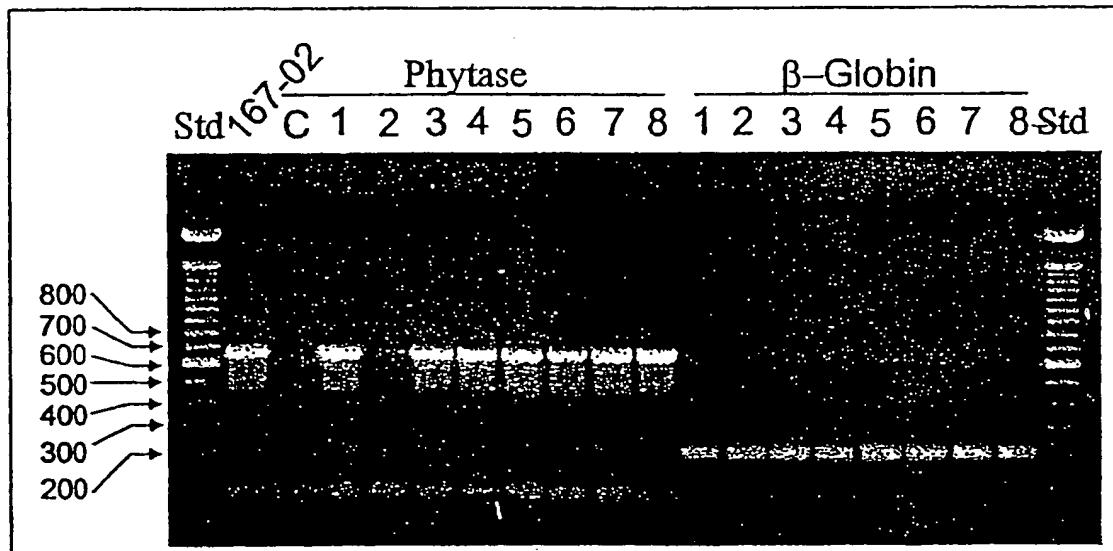


Figure 14

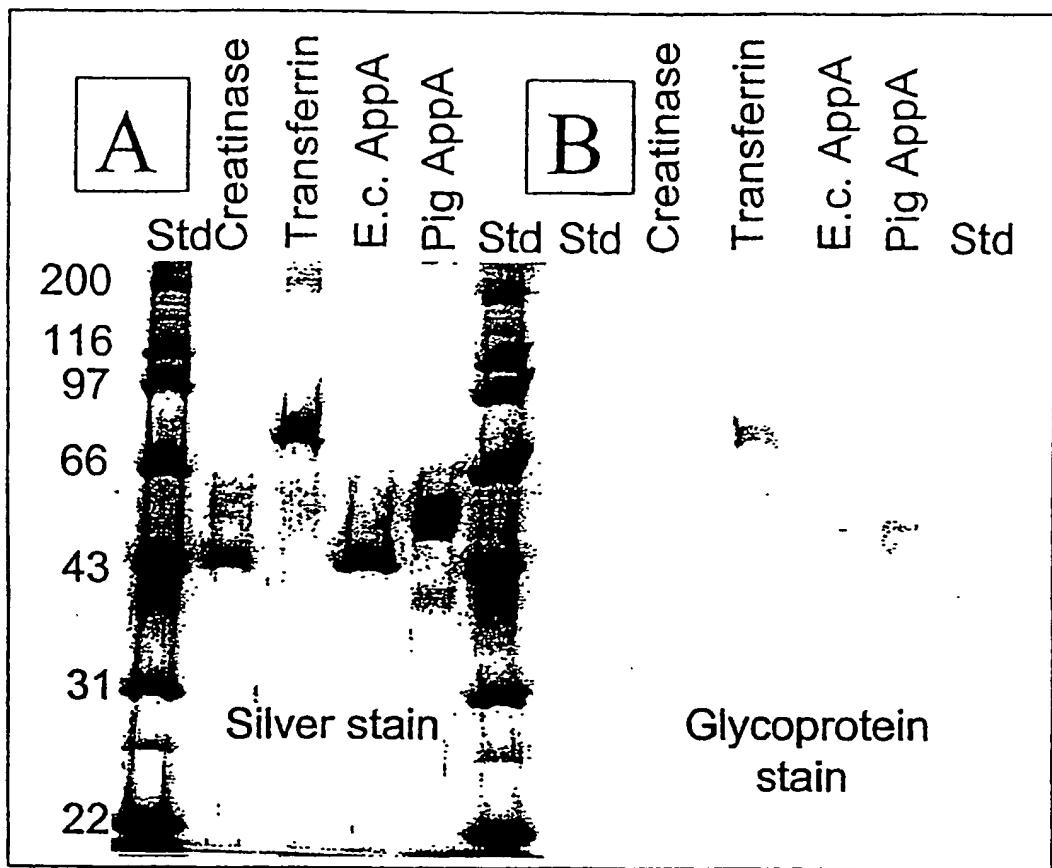


Figure 15

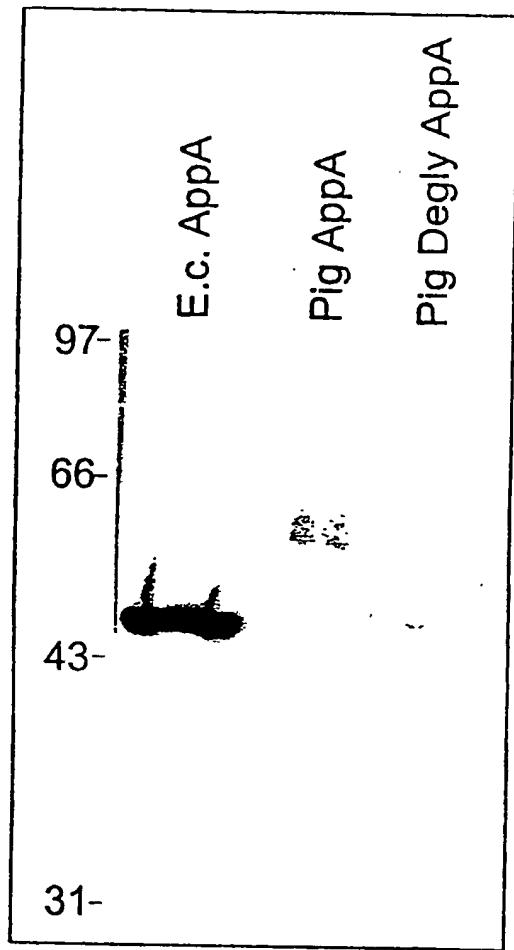


Figure 15B

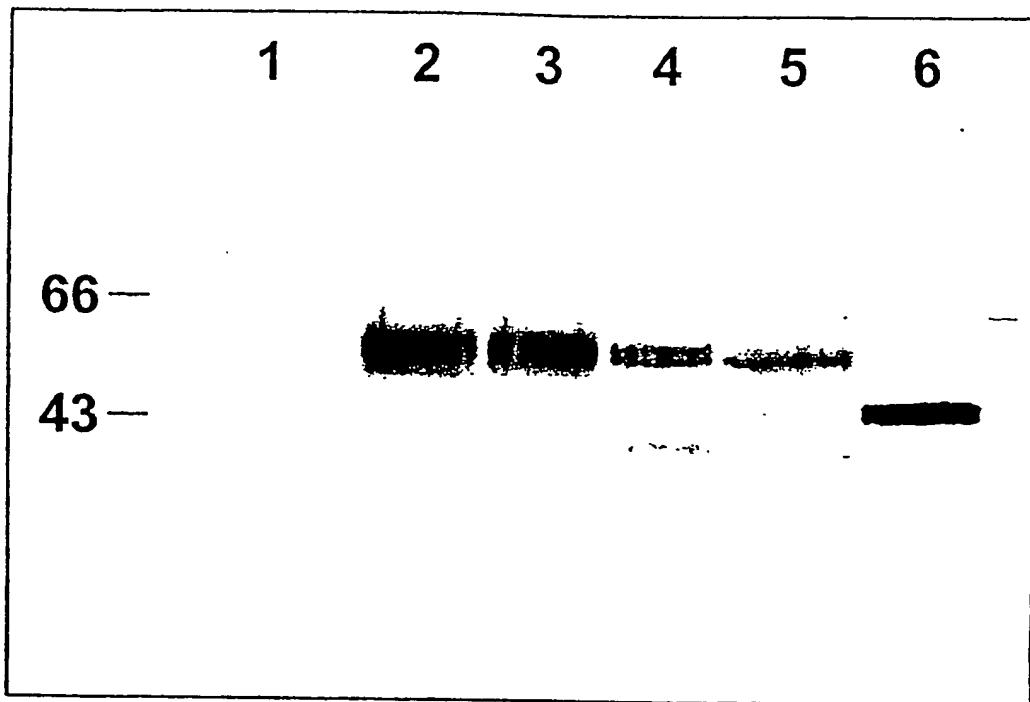


Figure 16

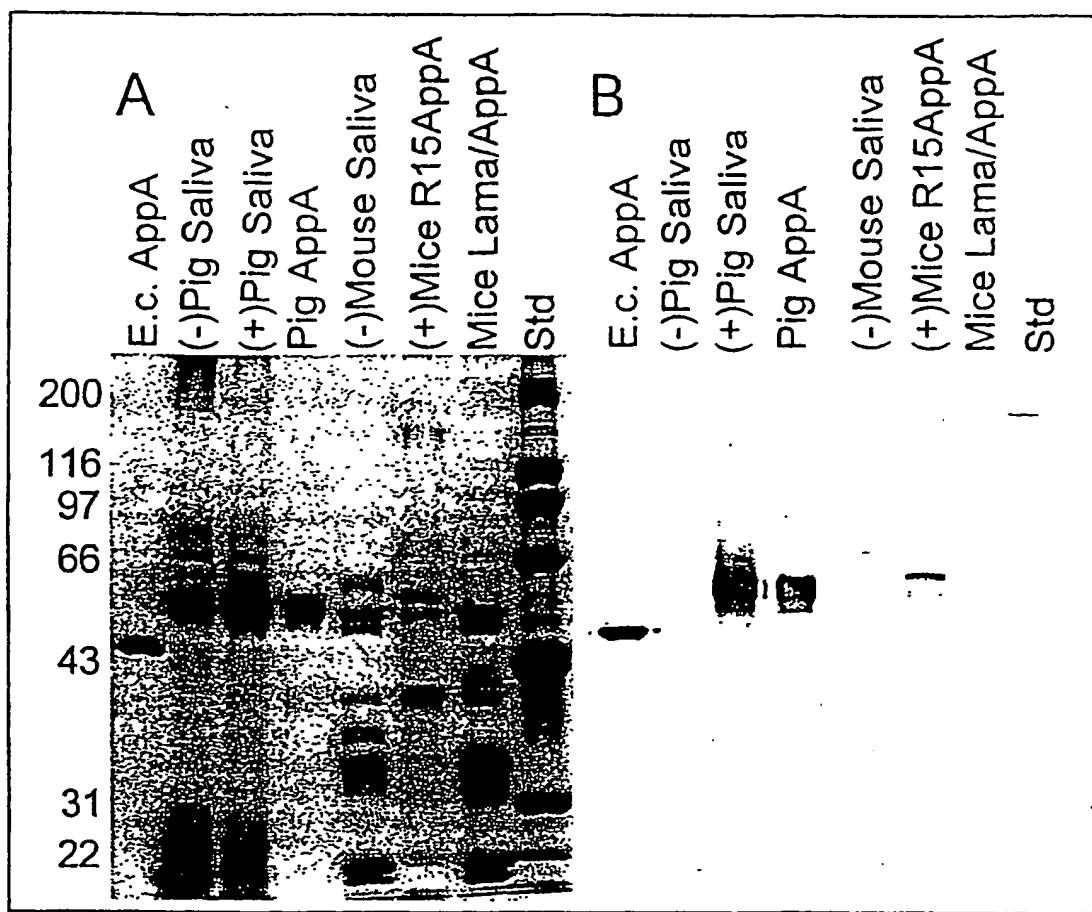


Figure 17

Figure 18: Nucleic acid sequence of the known segment of the R15/appa+intron plasmid, including the vector sequences of pBLCAT3 (SEQ ID NO:2).

LOCUS R15/appa+intron 6708 bp DNA SYN 15-APR-2000
 DEFINITION R15/appa+intron transgene with vector cut 13543 to 4954
 ACCESSION R15/appa+intron
 REFERENCE 1 (bases 1 to 6708))
 SOURCE synthetic construct.
 ORGANISM synthetic construct
 artificial sequence.
 KEYWORDS salivary proline-rich protein, acid glucose-1-phosphatase; appA gene; periplasmic phosphoanhydride phosphohydrolase; artificial sequence;
 AUTHORS Golovan, S., Forsberg, C.W., Phillips, J.
 JOURNAL Unpublished.

DEFINITION Rat salivary proline-rich protein (RP15) gene.
 ACCESSION M64793 M36414
 VERSION M64793.1 GI:206711
 SOURCE Rat (Sprague-Dawley) liver DNA.
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;
 Mammalia;
 Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 1748)
 AUTHORS Lin, H.H. and Ann, D.K.
 TITLE Molecular characterization of rat multigene family
 encoding proline-rich proteins
 JOURNAL Genomics 10, 102-113 (1991)
 MEDLINE 91257817
 FEATURES Location/Qualifiers
 source 1..1748
 /organism="Rattus norvegicus"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /tissue_type="liver"
 /tissue_lib="cosmid genomic library"
 misc_feature 1802-1810
 /function=" consensus sequence for initiation in
 higher eukaryotes "

FEATURES Location/Qualifiers
 DEFINITION E. coli periplasmic phosphoanhydride phosphohydrolase (appA) gene,
 ACCESSION M58708 L03370 L03371 L03372 L03373 L03374 L03375
 VERSION M58708.1 GI:145283
 SOURCE Escherichia coli DNA.
 ORGANISM Escherichia coli
 Bacteria; Proteobacteria; gamma subdivision;
 Enterobacteriaceae;
 Escherichia.
 REFERENCE 1 (bases 1811..3109)
 AUTHORS Dassa, J., Marck, C. and Boquet, P.L.

Figure 18 (continued):

TITLE The complete nucleotide sequence of the *Escherichia coli* gene *appA* reveals significant homology between pH 2.5 acid phosphatase and glucose-1-phosphatase
 JOURNAL *J. Bacteriol.* 172 (9), 5497-5500 (1990)
 MEDLINE 90368616

FEATURES **Location/Qualifiers**
 Source 1811..3109
 /organism="Escherichia coli"
 /db_xref="taxon:562"
 sig_peptide 1811.. 1876
 /gene="appA"
 CDS 1811..3109
 /gene="appA"
 /standard_name="acid phosphatase/phytase"
 /transl_table=11
 /product="periplasmic phosphoanhydride phosphohydrolase"
 /protein_id="AAA72086.1"
 /db_xref="GI:145285"

/translation="MKAILIPFLSLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAP
 TKATQLMQDVTPTDAWPTWPVKLGWLTPRGELIAYLGHYQRQRLVADGLLAKKGCPQS
 GQVAlIADVDERTRKTGEAFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNA
 NVTDAILSRAGGSIAADFTGHRQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPS
 ELKVSADNVSLTGAVSLASMLTEIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQF
 YLLQRTPEVARSRATPLLDLIKTALTPHPPQKQAYGVTLPSTSVLFIAGHDTNLANLGG
 ALELNWTLPGQPDNTPPGELVFERWRRILSDNSQWIQVSLVFPQTLQQMRDKTPLSLNT
 PPGEVKLTLAGCEERNAQGMCSLAGFTQIVNEARIPACSL"

mat_peptide 1877 3106
 /gene="appA"
 /product="periplasmic phosphoanhydride phosphohydrolase"

mutation replace(1817.. 1819, "gcg changed to gcc")
 /gene="appA"
 /standard_name="A3 mutant"
 /note="created by site directed mutagenesis"
 /phenotype="silent mutation"
 mutation replace(3092..3094, " ccg changed to ccc")
 /gene="appA"
 /standard_name=" P428 mutant"
 /note="created by site directed mutagenesis"
 /phenotype=" silent mutation "
 mutation replace(3095..3097, " gcg changed to gct")
 /gene="appA"
 /standard_name=" A429 mutant"
 /note="created by site directed mutagenesis"
 /phenotype=" silent mutation "

Figure 18 (continued):

DEFINITION Plasmid pBLCAT3 (bases 3109 to 6708)

ACCESSION X64409

VERSION X64409.1 GI:58163

SOURCE synthetic construct.

ORGANISM synthetic construct

artificial sequence.

REFERENCE 1 (bases 3109 to 6708)

AUTHORS Luckow, B.H.R.

TITLE Direct Submission

JOURNAL Submitted (06-FEB-1992) B.H.R. Luckow, German Cancer Res Center, Im Neuenheimer Feld 280, W-6900 Heidelberg, FRG

REFERENCE 2 (bases 3109 to 6708)

AUTHORS Luckow, B. and Schutz, G.

TITLE CAT constructions with multiple unique restriction sites

for the functional analysis of eukaryotic promoters and

regulatory elements

JOURNAL Nucleic Acids Res. 15 (13), 5490 (1987)

MEDLINE 87260024

COMMENT Promoterless CAT vector for transient transfection experiments with eukaryotic cells. Allows the analysis of foreign promoters and enhancers.

FEATURES

source Location/Qualifiers

3109 to 6116

/organism="synthetic construct"

/db_xref="taxon:32630"

SV40 t intron 3197..3810

/note="SV40 signals"

polyA_signal 3807..4047

/note="SV40 signals"

CDS complement(5244..6104)

/codon_start=1

/transl_table=11

/gene="Amp"

/product="beta-lactamase"

/protein_id="CAA45753.1"

/db_xref="GI:58165"

BASE COUNT 1916 a 1479 c 1515 g 1798 t

ORIGIN

1 GGATCCCTT TGCTATGTAG TTTTAATGG AAATTACAAC CCATAGTGTG TTGATAAATA

61 GAGAGTCCTG TTGGTTTAA GCAACCTCTG TTTCTCATAA ACTCCATAAA AACAGGAATA

121 CTCTTTGTTT CTAGCATAAC CAAAAGATT AGTGAATTGA AAACAATGTT CCCTTAGAGT

181 ATAGGTCTAA TAAACCCGAA AATATTACCA TGATTAATGAG CATTGTAGAG TATCTCATAG

241 CATGTAGTAT CCATAGTCCA TCAATGAGAG AGACATTTAA CATGATTTTC ATTAATCAGG

301 TGGAAAAGAC ATGACAACAT TCACAGGCAC TGCAAGAAC ATAGTGGTCC ACCTTGCACA

361 TATTTCACTA AACTAGGTTT ATCTATTTTG TTGCTTTCTC TAACATCTCT GCAATGAAGC

421 AGGTCAACAG TGCCACATAT CCTTTACTTA ACCTTAAGGAA CACAAAAAAAT TTCTACATA

481 TATCCTGGTT AGAGAGTGCT TAAAATAAGT TTCCAAGAA TGGAAAAGAA ATGTTCTGAC

541 TTAACAAATTA AGACAGTATT TATTTAAAGC AAGAAATATG AGGCACACAA GAAAATATTT

601 TGGGAAGAAA CCATTTGGTG AACAATATTT CAAATAAAAA TAGACAAACA TAGTTAATTG

661 TAAAACATAT GTTGACCAAG CCCTTCTTTT CAATAGGCTT AATGTGAATA AAATGTTAAA

721 GATTCTCTTT GGGTGGCTGC AAATGTCCA CGAATAAGAC AAAATATAAA AATAAGGACT

781 GAGTCTCACA AAATGAAAAG GAAATATATT CAGAAAGAGA ATCTTGAGAG AATGTGTTGT

841 CACAAATTAA AGAAAACCTG TGGTGAATGA CATCTGAGG CCTGAGCTAT TACTGACATT

Figure 18 (continued):

901 TAAGATAAAAG GTAACGTAT ACATTTGTCC CATTGAGGGG ACAAGAAAGC TGCTCTCATG
 961 TTCAGCTCTA TAATTCTTGC CTTAAACAAC TTAATAGAA TGATTTAAAAT TATGGAGCTG
 1021 TCCATGGACC TTTGAAATAT AAAATAGTCA AGCAACTTAT CAAGGAATTA CAGATTCCTT
 1081 GATACTAACA CAGGTAAATC CCACACGTGT TTTGAGACTA CATTGCTGG GATTTTATTG
 1141 ATGTAATAGG TCACATGTTT TTCGGGCCAA TGTTGCTGTT ATTGGTAC TTCAAGAGAA
 1201 TAGTGGCAAC TGATGCTATG TATTCTAGGG GTTGAAGTG ATGTTTCATG ATTGAAATTT
 1261 GTAAAAGAAT AACATCATCA TTCTTAACAA TAGAACATAT AAAGTCACAC AGAAGTGACA
 1321 GTGTTAACG TGTAATTG ATCAAAGAAA TTATACCT TCAGTTCAA TGGAAATAAT
 1381 TACTGATAAT ACAAACATGT GTGAACACAC ACTAATCCTA TCCAAATGCA CAGTGATACA
 1441 CAGAAAATAT TAGCAAGTAG AATGCAATAT TTATATAACG ATTGTATTAA TCAATCAATT
 1501 GTATGTATCA ATATATGGGC TATTTTCTTA CACATGATT TATTCAAAT TACTCTAATC
 1561 ATTGTTGAAC CATTAGAAA AGGCATACTG GCAACTTTTC CTTACCTCAT CCAGCTGGGC
 1621 AAAAGTCCC GTGTTGAGTA AAGGATGCAA GATTTCCTGC TCTGTTAAGT ATAAAATAAT
 1681 AGTATGAATT CAAAGGTGCC ATTCTTCTGC TTCTAGTTAT AAAGGCAGTG CTIGCTTCTT
 1741 CCAGCACAGA TCTGGATCTC GAGGAGCTG GCGAGATTTC CAGGAGCTAA GGAAGCTAAA
 1801 AGCCGCCACC ATGAAAGCCA TCTTAATCCC ATTTTATCT CTTCTGATTG CGTTAACCCC
 1861 GCAATCTGCA TTCGCTCAGA GTGAGCCCGA GCTGAAGCTG GAAAGTGTGG TGATTGTCAG
 1921 TCGTCATGGT GTGGTGCCTC CAACCAAGGC CACGCAACTG ATGCAGGATG TCACCCAGA
 1981 CGCATGGCCA ACCTGGCCGG TAAAACCTGGG TTGGTGCACA CCGCGCGGTG GTGAGCTAAT
 2041 CGCCTATCTC GGACATTACC AACGCCAGCG TCTGGTAGCC GACGGATTGC TGGCGAAAAA
 2101 GGGCTGCCCG CAGTCTGGTC AGGTCGCGAT TATTGCTGAT GTCGACGAGC GTACCCGTAA
 2161 AACAGGCGAA GCCTTCGCCG CCGGGCTGGC ACCTGACTGT GCAATAACCG TACATACCA
 2221 GGCAGATACG TCCAGTCCCG ATCCGTTATT TAATCCTCTA AAAACTGGCG TTTGCCAACT
 2281 GGATAACCGC AACGTGACTG ACGCCATCCT CAGCAGGGCA GGAGGGTCAA TTGCTGACTT
 2341 TACCGGGCAT CGGAAACCGG CGTTTCGCGA ACTGGAACGG GTGCTTAATT TTCCGCAATC
 2401 AAAACTGTGCA TTAAACGTG AGAAACAGGA CGAAAGCTGT TCATTAACGC AGGCAATTAC
 2461 ATCGGAACTC AAGGTGAGCG CCGACAATGT CTCACTTAAACC GGTGCGGTAA GCCTCGCATC
 2521 AATGCTGACG GAGATATTTC TCCCTGAAACA AGCACAGGGG ATGCCGGAGC CGGGGTGGGG
 2581 AAGGATCACC GATTACACCC AGTGGAACAC CTTGCTAAGT TTGCTATAACG CGCAATTTA
 2641 TTGCTACAA CGCACGCCAG AGGTTGCCCG CAGCCGCGCC ACCCCGTAT TAGATTGAT
 2701 CAAGACAGCG TTGACGCCCG ATCCACCCCA AAAACAGGCG TATGGTGTGA CATTACCCAC
 2761 TTCAGTGTG TTTATGCCG GACACGATAC TAATCTGGCA AATCTCGCG GCGCACTGGA
 2821 GCTCAACTGG ACCGTTCCCG GTCAGCCCGA TAACACGCCG CCAGGTGGTG AACTGGTGT
 2881 TGAACCGTGG CGTCCGGCTAA GCGATAACAG CCAGTGGATT CAGGTTTCGC TGTTCTCCA
 2941 GACTTTACAG CAGATGCGTG ATAAAACGCC GCTGCTTAA AATACGCCG CCGGAGAGGT
 3001 GAAACTGACC CTGGCAGGAT GTGAAGAGCG AAATGCGCAG GGCATGTGT CGTTGGCAGG
 3061 TTTTACGCAA ATCGTGAATG AAGCACCGAT ACCCGCTTGC AGTTTGTAAAG GTATAAGGCA
 3121 GTTATTGTTG CCCTTAAACG CCTGGTGCATA CGCCTGAAATA AGTATAAAATA AGGGATGAA
 3181 TGGCAGAAAAT TCGCCGGATC TTGTTGAAAG AACCTTACTT CTGTTGTTG ACATAATTGG
 3241 ACAAACTTAC TACAGAGATT TAAAGCTCTA AGGTTAAATAT AAAATTTTA AGTGTATAAT
 3301 GTGTTAAACT ACTGATTCTA ATTGTTGTTG TATTTTAGAT TCCAACCTAT GGAACGTGATG
 3361 AATGGGGAGCA GTGGTGGAAAT GCCCTTAAATG AGGAAAACCT GTTTTGCTCA GAAGAAATGC
 3421 CATCTAGTGA TGATGAGGCT ACTGCTGACT CTCAACATTC TACTCCTCCA AAAAAAGAAGA
 3481 GAAAGGTAGA AGACCCCAAG GACTTTCCCT CAGAATTGCT AAGTTTTTG AGTCATGCTG
 3541 TGTTTAGTAA TAGAACTCTT GCTTGTGTTG CTATTACAC CACAAAGGAA AAAGCTGCAC
 3601 TGCTATACAA GAAAATTATG AAAAATATT CTGTAACCTT TATAAGTAGG CATAACAGTT
 3661 ATAATCATAA CATACTGTTT TTTCTTACTC CACACAGGCA TAGAGTGTCT GCTATTAATA
 3721 ACTATGCTCA AAAATTGTGT ACCTTTAGCT TTTTAATTG TAAAGGGTTT AATAAGGAAT
 3781 ATTGATGTA TAGTGCCTTG ACTAGAGATC ATAATCAGCC ATACCAACATT TGTAGAGGTT
 3841 TTACTTGTCT TAAAAAAACCT CCCACACCTC CCCCTGAAAC TGAAACATAA AATGAATGCA
 3901 ATTGTTGTG TTAACGTGTT TATTGCAAGCT TATAATGGTT ACAAAATAAG CAATAGCATC
 3961 ACAAAATTCA CAAATAAAAGC ATTTTTTCA CTGCATTCTA GTTGTGGTTT GTCCAAACTC
 4021 ATCAATGTAT CTTATCATGT CTGGATCGAT CCCGGGGTAC CGAGCTCGAA TTCGTAATCA
 4081 TGGTCATAGC TGTTCCTGT GTGAATTGTT TATCCGCTCA CAATTCCACA CAACATACGA
 4141 GCCGGAAGCA TAAAGTGTAA AGCCTGGGT GCCTAATGAG TGAGCTAATC CACATTAATT
 4201 GCGTTGCGCT CACTGCCCGC TTTCCAGTGT GGAAACCTGT CGTGCAGCT GCATTAATGA
 4261 ATCGGGCAAC GCGCGGGGAG AGGCCTGGTTG CGTATTGGGC GCTCTCGCTC TTCCCTCGCTC
 4321 ACTGACTCGC TGCGCTCGGT CGTCCGGCTG CGGGGAGCGG TATCAGCTCA CTCAAAGGGC

Figure 18 (continued):

4381 GTAATACGGT TATCCACAGA ATCAGGGAT AACGCAGGAA AGAACATGTG AGCAAAAGGC
 4441 CAGCAAAAGG CCAGGAACCG TAAAAAGGCC GCGTTGCTGG CGTTTTCCA TAGGCTCCGC
 4501 CCCCCCTGACG AGCATCACAA AAATCGACGC TCAAGTCAGA GGTGGCGAAA CCCGACAGGA
 4561 CTATAAAGAT ACCAGGCAGTT TCCCCCTGGA AGCTCCCTCG TGCGCTCTCC TGTTCCGACC
 4621 CTGCCGCTTA CCGGATAACCT GTCCGCCTT CTCCCTCGG GAAGCGTGGC GCTTCTCAA
 4681 TGCTCACCGCT GTAGGTATCT CAGITCGGTG TAGGTGTTGC GCTCCAAGCT GGGCTGTGTG
 4741 CACGAACCCC CGGTTCAGCC CGACCGCTGC GCCTTATCCG GTAACATATCG TCTTGAGTCC
 4801 AACCCGGTAA GACACGACTT ATCGCCACTG GCAGCAGCCA CTGGTAACAG GATTAGCAGA
 4861 GCGAGGTATG TAGGCGGTGC TACAGAGTT TTGAAGTGGT GGCCTAACTA CGGCTACACT
 4921 AGAAGGACAG TATTTGGTAT CTGCGCTCG CTGAAGCCAG TTACCTTCGG AAAAGAGTT
 4981 GGTAGCTCTT GATCCGGCAA ACAAAACCACC GCTGGTAGCG GTGGTTTTTG TGTTTGCAAG
 5041 CAGCAGATTA CGCGCAGAAA AAAAGGATCT CAAGAAGATC CTTTGATCTT TTCTACGGGG
 5101 TCTGACGCTC AGTGGAACGA AAACTCACGT TAAGGGATTT TGGTCATGAG ATTATCAAAA
 5161 AGGATCTTCA CCTAGATCCT TTAAATTAA AAATGAAGTT TTAAATCAAT CTAAAGTATA
 5221 TATGAGTAAA CTTGGTCTGA CAGTTACCAA TGCTTAATCA GTGAGGCACC TATCTCAGCG
 5281 ATCTGTCTAT TTGTTCATC CATAGTTGCC TGACTCCCCG TCGTGTAGAT AACTACGATA
 5341 CGGGAGGGCT TACCATCTGG CCCCAGTGCT GCAATGATAC CGCGAGACCC ACGCTCACCG
 5401 GCTCCAGATT TATCAGCAAT AAACCGACCA GCCGGAAGGG CCGAGCGCAG AAGTGGTCCCT
 5461 GCAACTTAT CCGCCTCCAT CCAGTCTATT AATTGTTGCC GGGAAAGCTAG AGTAAGTAGT
 5521 TCGCCAGTTA ATAGTTGCC CAACGTTGTT GCCATTGCTA CAGGCATCGT GGTGTCACGC
 5581 TCGTCGTTTG GTATGGCTTC ATTCAAGCTCC GGTTCCCAAC GATCAAGGGG AGTTACATGA
 5641 TCCCCCATGT TGTCGAAAAA AGCGGTTAGC TCCTTCGGTC CTCCGATCGT TGTCAGAAGT
 5701 AAGTTGGCCG CAGTGTATC ACTCATGGTT ATGGCAGCAC TGCATAATTTC TCTTACTGTC
 5761 ATGCCATCCG TAAGATGCTT TTCTGTGACT GGTGAGTACT CAACCAAGTC ATTCTGAGAA
 5821 TAGTGTATGC GGGCACCGAG TTGCTCTTGC CCGGCGTCAA TACGGGATAAA TACCGCGCCA
 5881 CATAGCAGAA CTTTAAAAGT GCTCATCAIT GGAAACAGTT CTICGGGGCG AAAACTCTCA
 5941 AGGATCTTAC CGCTGTGAG ATCCAGTCG ATGTAACCCCA CTCGTGACCC CAACTGATCT
 6001 TCAGCATCTT TTACTTTAC CAGCCTTCT GGGTGGACAA AAACAGGAAG GCAAAATGCC
 6061 GCAAAAAGG GAATAAGGGC GACACGGAAA TGTTGAATAC TCATACTCTT CCTTTTTCAA
 6121 TATTATTGAA GCATTATCA GGTTTATTGT CTCAATGAGCG GATACATATT TGAATGTATT
 6181 TAGAAAAATA AACAAATAGG GGTTCCGGCG ACATTCCCC GAAAAGTGCC ACCTGACGTC
 6241 TAAGAAAACCA TTATTATCAT GACATTAACC TATAAAAATA GGCGTATCAC GAGGCCCTTT
 6301 CGTCTCGCGC GTTTCGGTGA TGACGGTGA AACCTCTGAC ACATGCAGCT CCCGGAGACG
 6361 GTCACAGCTT GTCTGTAAAGC GGATGCCGGG AGCAGACAAG CCCGTCAAGG CGCGTCAAGCG
 6421 GGTGTTGGCG GGTGTCGGGG CTGGCTTAAC TATGCGGCAT CAGAGCAGAT TGTACTGAGA
 6481 GTGCACCCATA TGCGGTGTGA AATACCGCAC AGATGCGTAA GGAGAAAATA CGGCATCAGG
 6541 CGCCATTCGC CATTCAAGGCT GCGCAACTGT TGGGAAGGGC GATCGGTGGC GGCCTCTCG
 6601 CTATTACGCC AGCTGGCGAA AGGGGGATGT GCTGCAAGGC GATTAAGTTG GGTAACGCCA
 6661 GGGTTTTCCC AGTCACGACG TTGTAAAACG ACGGCCAGTG CCAAGCTT

//

Figure 19: Nucleic acid sequence of the known segment of the R15/appa+intron transgene used for the generation of transgenic mice (SEQ ID NO: 3).

LOCUS R15/appa 4060 bp DNA SYN 15-APR-2000
 DEFINITION R15/appa transgene without vector
 ACCESSION R15/appa
 REFERENCE 1 (bases 1 to 4060)
 SOURCE synthetic construct.
 ORGANISM synthetic construct
 artificial sequence.
 KEYWORDS salivary proline-rich protein, acid glucose-1-phosphatase; appA
 gene; periplasmic phosphoanhydride phosphohydrolase; artificial
 sequence;
 AUTHORS Golovan, S., Forsberg, C.W., Phillips, J.
 JOURNAL Unpublished.

DEFINITION Rat salivary proline-rich protein (RP15) gene.
 ACCESSION M64793 M36414
 VERSION M64793.1 GI:206711
 SOURCE Rat (Sprague-Dawley) liver DNA.
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;
 Mammalia;
 Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 1748)
 AUTHORS Lin, H.H. and Ann, D.K.
 TITLE Molecular characterization of rat multigene family
 encoding proline-rich proteins
 JOURNAL Genomics 10, 102-113 (1991)
 MEDLINE 91257817
 FEATURES Location/Qualifiers
 source 1..1748
 /organism="Rattus norvegicus"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /tissue_type="liver"
 /tissue_lib="cosmid genomic library"
 misc_feature 1802-1810
 /function=" consensus sequence for initiation in
 higher eukaryotes "

FEATURES Location/Qualifiers
 DEFINITION E. coli periplasmic phosphoanhydride phosphohydrolase (appA)
 gene,
 ACCESSION M58708 L03370 L03371 L03372 L03373 L03374 L03375
 VERSION M58708.1 GI:145283
 SOURCE Escherichia coli DNA.
 ORGANISM Escherichia coli
 Bacteria; Proteobacteria; gamma subdivision;
 Enterobacteriaceae;
 Escherichia.
 REFERENCE 1 (bases 1811..3109)
 AUTHORS Dassa, J., Marck, C. and Boquet, P.L.

Figure 19 (continued):

TITLE The complete nucleotide sequence of the *Escherichia coli* gene *appA* reveals significant homology between pH 2.5 acid phosphatase and glucose-1-phosphatase
 JOURNAL *J. Bacteriol.* 172 (9), 5497-5500 (1990)
 MEDLINE 90368616

FEATURES **Location/Qualifiers**
 Source 1811..3109
 /organism="Escherichia coli"
 /db_xref="taxon:562"
 sig_peptide 1811..1876
 /gene="appA"
 CDS 1811..3109
 /gene="appA"
 /standard_name="acid phosphatase/phytase"
 /transl_table=11
 /product="periplasmic phosphoanhydride phosphohydrolase"
 /protein_id="AAA72086.1"
 /db_xref="GI:145285"

/translation="MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAP
TKATQLMQDVTPDAWPTWPVKLGWLTPRGELIAYLGHYQRQRLVADGLLAKKGCPQS
GQVAIIDVDERTRKTEAFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNA
NVTDAILSRAGGSIADFTGHRQTAFRELERVLFNFPQSNLCLKREKQDESCSLTQALPS
ELKVSADNVSLTGAVSLASMLTEIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQF
YLLQRTPEVARSRATPLLDLIKTALTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGG
ALELNWTLPQPDNTPPGGELVPERWRRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNT
PPGEVKLTLAGCEERNAQGMCSLAGFTQIVNEARIPACSL"

mat_peptide 1877 3106
 /gene="appA"
 /product="periplasmic phosphoanhydride phosphohydrolase"

mutation replace(1817.. 1819, "gcg changed to gcc")
 /gene="appA"
 /standard_name="A3 mutant"
 /note="created by site directed mutagenesis"
 /phenotype="silent mutation"
 mutation replace(3092..3094, " ccg changed to ccc")
 /gene="appA"
 /standard_name=" P428 mutant"
 /note="created by site directed mutagenesis"
 /phenotype=" silent mutation "
 mutation replace(3095..3097, " gcg changed to gct")
 /gene="appA"
 /standard_name=" A429 mutant"
 /note="created by site directed mutagenesis"
 /phenotype=" silent mutation "

Figure 19 (continued):

```

SV40 t intron 3197..3810
                /note="SV40 signals"
polyA_signal    3807..4047
                /note="SV40 signals"

```

```

BASE COUNT      1257 a      814 c      843 g      1146 t
ORIGIN

```

```

1 GGATCCCCCT TGCTATGTAG TTTTAATGG AAATTACAAC CCATAGTGTG TTGATAAATA
61 GAGAGTCCTG TTTGTTTAA GCAACCTCTG TTTCTCATAA ACTCCATAAA AACAGGAATA
121 CTCTTGTCTT CTAGCATAAC CAAAGATTT AGTGAATTGA AAACAATGTT CCCTTAGAGT
181 ATAGGTCTAA TAACCCGAA ATATTACCA TGATACTGAG CATTGTAAAG TATCTCATAG
241 CATGTAGTAT CCATAGTCCA TCAATGAGAG AGACATTAA CATGATTTC ATTAAATCAGG
301 TGGAAAAGAC ATGACAACAT TCACAGGCAC TGACACAGAAC ATAGTGGTCC ACCTTGACAC
361 TATTCACTA AACTAGGTTT ATCTATTGT TGCTTCTC TAACATCTCT GCAATGAAGC
421 AGGTCAACAG TGCCACATAT CCTTTACTTA ACCTAAGGAA CACAAAAAAAT TTCTACATA
481 TATCCTGGTT AGAGAGTGT TAAAATAAGT TTCTTAAAGAA AGGACACACAA GAAAATATTT
541 TTAACAATTA AGACAGTATT TATTTAAAGC AAGAAATATG AGGCACACAA GAAAATATTT
601 TGGGAAGAAA CCATTGGTG AACAAATATT CAAATAAAA TAGACAAACA TAGTTAATTG
661 TAAAACATAT GTTGACCAG CCCTTCTT CAATAGGCTT ATGTGAATA AAATGTAAA
721 GATTCTCTT GGGTGGCTGC AAATTGTCCA CGAATAAGAC AAATATATAAA AATAAGGACT
781 GAGTCTCACA AAATGAAAAG GAAATATATT CAGAAAGAGA ATCTTGAGAG AATGTGTTGT
841 CACAAATTAA AGAAAACCTG TGGTGAATGA CATCCTGAGG CCTGAGCTAT TACTGACATT
901 TAAGATAAAG GTAATGTAT ACATTTGTCC CATTGAGGGG ACAAGAAAGC TGCTCTCATG
961 TTCAGCTCTA TAATTCTTGC CTTAAACAAAC TTAAATAGAA TGATTTAAAAA TATGGAGCTG
1021 TCCATGGACC TTTGAAATAT AAAATAGTCA AGCAACATTAT CAAGGAATTAA CAGATTCTT
1081 GATACTAACCA CAGGTAAATC CCACACGTGT TTGAGACTA CATTGCTGG GATTTTATTG
1141 ATGTAATAGG TCACATGTTT TTGAGGCTT CTTGCTGTGTT ATTGGTACAC TTCAAGAGAA
1201 TAGTGGCAAC TGATGCTATG TATTCTAGGG TTGAGGTG ATGTTTCAATG ATTGAAATT
1261 GTAAAAGAAT AACATCATCA TTCTTAACAA TAGAACATAT AAAGTCACAC AGAAGTGACA
1321 GTGTTTAAGC TGTACTATTG ATCAAAGAAA TTATTACCT TCAGTTTCAA TGGAAATAAT
1381 TACTGATAAT ACAAAACATGT GTGAACACAC ACTAACTCTA TCCAAATGCA CAGTGATACA
1441 CAGAAAATAT TAGCAAGTAG AATGCAATAT TTATATAACG ATTGTATTAA TCAATCAATT
1501 GTATGTATCA ATATATGGG TATTCTCTA CACATGATT TATTCAAATT TACTCTAATC
1561 ATTGTTGAAC CATTAGAAA AGGCATACTG GCAACTTTTC CTTACCTCAT CCAGCTGGC
1621 AAAAGTCCCAGTGTGAGTA AAGGATGCAA GATTCTCTGC TCTGTTAAGT ATAAAATAAT
1681 AGTATGAATT CAAAGGTGCG ATTCTCTGC TTCTAGTTAT AAAGGCAGTG CTGCTTCTT
1741 CCAGCACAGA TCTGGATCTC GAGGAGCTG GCGAGATTTT CAGGAGCTAA GGAAGCTAAA
1801 AGCCGCCACC ATGAAAGCCA TCTTAATCCC ATTCTTATCT CTTCTGATTC CGTTAACCCC
1861 GCAATCTGCA TTGCTCAGA GTGAGCCGGA GCTGAAGCTG GAAAGTGTGG TGATTGTCAG
1921 TCGTATGGT GTGGTGTCTC CAACCAAGGC CACGCAACTG ATGCAGGATG TCACCCAGA
1981 CGCATGGCCA ACCTGGCGGG TAAAATGGG TTGGCTGACA CCGCGCGGTG GTGAGCTAAT
2041 CGCCTATCTC GGACATTACC AACGCCAGCG TCTGGTAGCC GACGGATTGC TGGCGAAAAAA
2101 GGGCTGCCCG CAGTCTGGTC AGGTGCGGAT TATTGCTGAT GTGACGAGC GTACCCGTAA
2161 AACAGGGAA GCTTCGCGG CCGGGCTGGC ACCTGACTGT GCAATAACCG TACATACCA
2221 GGCAGATAACG TCCAGTCCCCG ATCCGTATT TAATCTCTA AAAACTGGCG TTGCAACT
2281 GGATAACGCG AACGTGACTG ACGCGATCTC CAGCAGGGCA GGAGGGTCAA TTGCTGACTT
2341 TACCGGGCAT CGGAAACGG CGTTCTGGGA ACTGGAACGG GTGCTTAATT TTCCGCAATC
2401 AAACTTGTGC TTAAACGTG AGAAACAGGA CGAAAGCTGT TCATTAACCG AGGCATTACC
2461 ATCGGAACCTC AAGGTGAGCG CCGACAATGT CTCATTAACC GGTGCGGTAA GCCTCGCATC
2521 AATGCTGACG GAGATATTTC TCCCTGAAACA AGCACAGGGAA ATGCCGGAGC CGGGGTGGGG
2581 AAGGATCACC GATTCACACC AGTGGAACAC CTTGCTAAGT TTGCTAATCG CGCAATTAA
2641 TTGCTACAA CGCACGCCAG AGGTGCGCG CAGCCGCGCC ACCCCGTTAT TAGATTGAT
2701 CAAGACAGCG TTGACGCCCC ATCCACCGCA AAAACAGGCG TATGGTGTGA CATTACCCAC
2761 TTGCTGCTG TTATCGCCG GACACGATAC TAATCTGGCA ATCTCGGGC GCGCACTGGA
2821 GCTCAACTGG ACGCTTCCCCG GTCAAGCCGA TAACACGCCG CCAGGTGGTG AACTGGTGT
2881 TGAACGCTGG CGTCGGCTAA GCGATAACAG CCAGTGGATT CAGGTTTCGC TGGTCTTCCA

```

Figure 19 (continued):

2941 GACTTTACAG CAGATGCGTG ATAAAAGGCC GCTGTCATTA AATACGCCGC CCGGAGAGGT
3001 GAAACTGACC CTGGCAGGAT GTGAAGAGCG AAATGCGCAG GGCATGTGTT CGTTGGCAGG
3061 TTTTACGCCAA ATCGTGAATG AAGCACGCAT ACCCGCTTGC AGTTTGTAAAG GTATAAGGCA
3121 GTTATTGGTG CCCTTAAACG CCTGGTGCTA CGCCTGAATA AGTGATAATA AGCGGATGAA
3181 TGGCAGAAAT TCGCCGGATC TTTGTGAAGG AACCTTACTT CTGTGGTGTG ACATAATTGG
3241 ACAAACTACC TACAGAGATT TAAAGCTCTA AGGTTAAATAT AAAATTTTTA AGTGTATAAT
3301 GTGTTAAACT ACTGATTCTA ATTGTTGTG TATTTTAGAT TCCAACTTAT GGAACTGATG
3361 AATGGGAGCA GTGGTGGAAAT GCCTTTAATG AGGAAAACCT GTTTTGCTCA GAAGAAATGC
3421 CATCTAGTGA TGATGAGGCT ACTGCTGACT CTCACACATTC TACTCCTCCA AAAAAAGAAGA
3481 GAAAGGTAGA AGACCCCAAG GACTTTCCCT CAGAATTGCT AAGTTTTTG AGTCATGCTG
3541 TGTTTAGTAA TAGAACTCTT GCTTGCTTGT CTATTACAC CACAAAGGAA AAAGCTGCAC
3601 TGCTATACAA GAAAATTATG GAAAATAATT CTGTAACCTT TATAAGTAGG CATAACAGTT
3661 ATAATCATAA CATACTGTTT TTTCTTACTC CACACAGGCA TAGAGTGCT GCTATTAATA
3721 ACTATGCTCA AAAATTGTGT ACCTTTAGCT TTTTAATTG TAAAGGGTT AATAAGGAAT
3781 ATTTGATGTA TAGTGCTTG ACTAGAGATC ATAATCAGCC ATACCCACATT TGTAGAGGTT
3841 TTACTTGCTT TAAAAAAACCT CCCACACCTC CCCCTGAACC TGAAACATAA AATGAATGCA
3901 ATTGTTGTTG TTAACTTGTT TATTGAGCT TATAATGGTT ACAAAATAAG CAATAGCATC
3961 ACAAAATTC ACAAATAAAGC ATTTTTTCA CTGCATTCTA GTTGTGGTTT GTCCAAACTC
4021 ATCAATGTAT CTTATCATGT CTGGATCGAT CCCCCGGTAC

//

Figure 20: Nucleic acid sequence of the known segment of the R15/appa plasmid (including the vector sequences of pBLCAT3 (SEQ ID NO:4).

LOCUS R15/appa 6116 bp DNA SYN 15-APR-2000
 DEFINITION R15/appa transgene with vector
 ACCESSION R15/appa
 REFERENCE 1 (bases 1 to 6116)
 SOURCE synthetic construct.
 ORGANISM synthetic construct
 artificial sequence.
 KEYWORDS salivary proline-rich protein, acid glucose-1-phosphatase; appA
 gene; periplasmic phosphoanhydride phosphohydrolase; artificial
 sequence;
 AUTHORS Golovan, S., Forsberg, C.W., Phillips, J.
 JOURNAL Unpublished.

DEFINITION Rat salivary proline-rich protein (RP15) gene.
 ACCESSION M64793 M36414
 VERSION M64793.1 GI:206711
 SOURCE Rat (Sprague-Dawley) liver DNA.
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;
 Mammalia;
 Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 1748)
 AUTHORS Lin, H.H. and Ann, D.K.
 TITLE Molecular characterization of rat multigene family
 encoding proline-rich proteins
 JOURNAL Genomics 10, 102-113 (1991)
 MEDLINE 91257817
 FEATURES Location/Qualifiers
 source 1..1748
 /organism="Rattus norvegicus"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /tissue_type="liver"
 /tissue_lib="cosmid genomic library"
 misc_feature 1802-1810
 /function=" consensus sequence for initiation in
 higher eukaryotes "

FEATURES Location/Qualifiers
 DEFINITION E. coli periplasmic phosphoanhydride phosphohydrolase (appA)
 gene,

ACCESSION MS8708 L03370 L03371 L03372 L03373 L03374 L03375
 VERSION MS8708.1 GI:145283
 SOURCE Escherichia coli DNA.
 ORGANISM Escherichia coli
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 Escherichia.

REFERENCE 1 (bases 1811..3109)
 AUTHORS Dassa, J., Marck, C. and Boquet, P.L.
 TITLE The complete nucleotide sequence of the Escherichia coli gene appA
 reveals significant homology between pH 2.5 acid phosphatase
 and glucose-1-phosphatase
 JOURNAL J. Bacteriol. 172 (9), 5497-5500 (1990)

Figure 20 (continued):

MEDLINE 90368616

FEATURES	Location/Qualifiers
Source	1811..3109 /organism="Escherichia coli" /db_xref="taxon:562"
sig_peptide /gene="appA"	1811..1876
CDS	1811..3109 /gene="appA" /standard_name="acid phosphatase/phytase" /transl_table=11 /product="periplasmic phosphoanhydride phosphohydrolase" /protein_id="AAA72086.1" /db_xref="GI:145285"
/translation="MKAILIPFLSLLIPLTPQSFAQSEPELKLESVVIVSRHGVRAP TKATQLMQDVTPTDAWPTWPVKGWLTPRGGEIAYLGHYQRQRILVADGLLAKKGCPQS GQVAIIDVDERTRKTGEAFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNA NVTDAILSRAGGSIADFTGHRQTAFRELERVINFPQSNLCLKREKQDESCSLTQALPS ELKVSADNVSLTGAVALASMLTEIFLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQF YLLQRTPEVARSRATPLLDLIKTALTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGG ALELNWTLPGQPDNTPPGGELVFERWRRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNT PPGEVKLTILAGCEERNAQGMCSLAGFTQIVNEARI PACSL"	
mat_peptide	1877 3106 /gene="appA" /product="periplasmic phosphoanhydride phosphohydrolase"
mutation	replace(1817.. 1819, "gcg changed to gcc") /gene="appA" /standard_name="A3 mutant" /note="created by site directed mutagenesis" /phenotype="silent mutation"
mutation	replace(3092..3094, " ccg changed to ccc") /gene="appA" /standard_name=" P428 mutant" /note="created by site directed mutagenesis" /phenotype=" silent mutation "
mutation	replace(3095..3097, " gcg changed to gct") /gene="appA" /standard_name=" A429 mutant" /note="created by site directed mutagenesis" /phenotype=" silent mutation "

DEFINITION Plasmid pBLCAT3 (bases 3109 to 6116)

ACCESSION	X64409
VERSION	X64409.1 GI:58163
SOURCE	synthetic construct.
ORGANISM	synthetic construct
	artificial sequence.
REFERENCE	1 (bases 3109 to 6116)
AUTHORS	Luckow, B.H.R.
TITLE	Direct Submission
JOURNAL	Submitted (06-FEB-1992) B.H.R. Luckow, German Cancer Res Center, Im Neuenheimer Feld 280, W-6900 Heidelberg, FRG

Figure 20 (continued):

REFERENCE 2 (bases 3109 to 6116)
 AUTHORS Luckow, B. and Schutz, G.
 TITLE CAT constructions with multiple unique restriction sites
 for the functional analysis of eukaryotic promoters and
 regulatory elements
 JOURNAL Nucleic Acids Res. 15 (13), 5490 (1987)
 MEDLINE 87260024
 COMMENT Promoterless CAT vector for transient transfection
 experiments with eukaryotic cells. Allows the analysis of foreign
 promoters and enhancers.
 FEATURES Location/Qualifiers
 source 3109 to 6116
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 polyA_signal 3262..3457
 /note="SV40 signals"
 CDS complement(4654..5514)
 /codon_start=1
 /transl_table=11
 /gene="Amp"
 /product="beta-lactamase"
 /protein_id="CAA45753.1"
 /db_xref="GI:58165"

BASE COUNT 1724 a 1386 c 1407 g 1599 t
 ORIGIN

1 GGATCCCCTT TGCTATGTAG TTTTTAATGG AAATTACAAC CCATAGTGTG TTGATAAATA
 61 GAGAGTCCTG TTGGTTTAA GCAACCTCTG TTTCTCATAA ACTCCATAAA AACAGGAATA
 121 CTCTTTGTTT CTAGCATAAC CAAAAGATT AGTGAATTGA AAACAATGTT CCCTTAGAGT
 181 ATAGGTCTAA TAACCCCGAA AATATTACCA TGATACTGAG CATTGTAAG TATCTCATAG
 241 CATGTAGTAT CCATAGTCCA TCAATGAGAG AGACATTAA CATGATTTTC ATTAATCAGG
 301 TGGAAAAGAC ATGACAACAT TCACAGGCAC TGCACAGAAC ATAGTGGTCC ACCTTGCACA
 361 TATTTCACTA AACTAGGTTT ATCTATTGTT TGCTTTCTC TAACATCTCT GCAATGAAGC
 421 AGGTCAACAG TGCCACATAT CCTTTACTTA ACCTAAGGAA CACAAAAAAT TTTCTACATA
 481 TATCCTGGTT AGAGAGTGT TAAAATAAGT TTTCCAAGAA TGGAAAAGAA ATGTTCTGAC
 541 TTAACAATTA AGACAGTATT TATTTAAAGC AAGAAATATG AGGCACACAA GAAAATATTT
 601 TGGGAAGAAA CCATTTGGTG AACAATATT CAAATAAAA TAGACAAACA TAGTTAATTG
 661 TAAAACATAT GTTGACCAG CCCTTCTTT CAATAGGCTT AATGTGAATA AAATGTTAAA
 721 GATTCTCTTT GGGTGGCTGC AAATTGTCCA CGAATAAGAC AAAATATAAA AATAAGGACT
 781 GAGTCTCACA AAATGAAAAG GAAATATATT CAGAAAGAGA ATCTTGAGAG AATGTGTTGT
 841 CACAAATTAA AGAAAACCTG TGGTGAATGA CATCTGAGG CCTGAGCTAT TACTGACATT
 901 TAAGATAAAG GTAATGTAT ACATTTGTCC CATTGAGGGG ACAAGAAAAGC TGCTCTCATG
 961 TTCAGCTCTA TAATTCTTGC CTTAAACAAAC TTAATAGAA TGATTTAAA TATGGAGCTG
 1021 TCCATGGACC TTGAAATAT AAAATAGTCA AGCAACTTAT CAAGGAATTAA CAGATTCTT
 1081 GATACTAACAA CAGGTAAATC CCACACGTGT TTGAGACTA CATTGCTGG GATTTATTG
 1141 ATGTAATAGG TCACATGTTT TTCGGGCCAA TGTGCTGTT ATTCGGTTAC TTCAAGAGAA
 1201 TAGTGGCAAC TGATGCTATG TATTCTAGGG GTTGAAGTG ATGTTTCATG ATTGAAATTT
 1261 GTAAAAGAAT AACATCATCA TTCTTAACAA TAGAACATAT AAAGTCACAC AGAAGTGACA
 1321 GTGTTAAGC TGTACTATTG ATCAAAGAAA TTATTTACCT TCAGTTCAA TGGAAATAAT
 1381 TACTGATAAT ACAAACATGT GTGAACACAC ACTAATCCTA TCCAAATGCA CAGTGATACA
 1441 CAGAAAATAT TAGCAAGTAG AATGCAATAT TTATATAACG ATTGTATTAA TCAATCAATT
 1501 GTATGTATCA ATATATGGGC TATTTCTTA CACATGATT TATTCAAATT TACTCTAATC
 1561 ATTGTTGAAC CATTAGAAA AGGCATACTG GCAACTTTTC CTTACCTCAT CCAGCTGGGC
 1621 AAAAGTCCCAGTGTGGAGTA AAGGATGCAA GATTCTCTGC TCTGTTAAGT ATAAAATAAT

Figure 20 (continued):

1681 AGTATGAATT CAAAGGTGCC ATTCTTCTGC TTCTAGTTAT AAAGGCAGTG CTTGCTTCTT
 1741 CCAGCACAGA TCTGGATCTC GAGGAGCTG GCGAGATTTT CAGGAGCTAA GGAAGCTAAA
 1801 AGCCGCCACC ATGAAAGCCA TCTTAATCCC ATTTTATCT CTTCTGATT CGTTAACCCC
 1861 GCAATCTGCA TTCGCTCAGA GTGAGCCGGA GCTGAAGCTG GAAAGTGTGG TGATTGTCAG
 1921 TCGTCATGGT GTGGTGTGCTC CAACCAAGGC CACGCAACTG ATGCAGGATG TCACCCCAGA
 1981 CGCATGGCCA ACCTGGCCGG TAAAATGGG TTGGCTGACA CGCGCCGGTG GTGAGCTAAT
 2041 CGCCTATCTC GGACATTACC AACGCCAGCG TCTGGTAGCC GACGGATTGC TGGCGAAAAA
 2101 GGGCTGCCCG CAGTCTGGTC AGGTGCGGAT TATTGCTGAT GTCGACGAGC GTACCCGTAA
 2161 AACAGGGAA GCCCTCGCCG CGGGGCTGGC ACCTGACTGT GCAATAACCG TACATACCCA
 2221 GGCAGATAACG TCCAGTCCCG ATCCGTIAATT TAATCCTCTA AAAACTGGGTTTGCCTAACT
 2281 GGATAACGCG AACGTGACTG ACGCGATCCT CAGCAGGGCA CGAGGGTCAA TTGCTGACTT
 2341 TACCGGGCAT CGGCAAACGG CGTTTCGCGA ACTGGAACGG GTGCTTAATT TTCCGCAATC
 2401 AAACTTGTGC CTAAACGTG AGAAACAGGA CGAAAGCTGT TCATTAACGC AGGCATTACC
 2461 ATCGGAACTC AAGGTGAGCG CGCACAATGT CTCATTAACCG GTGCGGGTAA GCCTCGCCTC
 2521 AATGCTGACG GAGATATTTT TCCTGCAACA AGCACAGGGAA ATGCCGGAGC CGGGGTGGGG
 2581 AAGGATCACC GATTACACC AGTGGAACAC CTTGCTAAGT TTGCTAACG CGCAATTITTA
 2641 TTGCTACAA CGCACGCCAG AGGTTGCGG CAGCCGCGCC ACCCGTTTAT TAGATTGAT
 2701 CAAGACAGCG TTGACGCCCG ATCCACCGCA AAAACAGGGC TATGGTGTGA CATTACCCAC
 2761 TTCAGTGTG TTTATCGCCG GACACGATAC TAATCTGGCA AATCTGGCG GCGCACTGG
 2821 GCTCAACTGG ACGCTTCCCG GTCAAGGGCA TAACACGCCG CCAGGTGGT AACTGGTGT
 2881 TGAACGCTGG CGTCCGCTAA GCGATAACAG CCAGTGGATT CAGGTTTCCG TGGTCTTCCA
 2941 GACTTTACAG CAGATGGTG ATAAAACCCC GCTGTCATTA ATAACGCCG CCGGAGAGGT
 3001 GAAACTGACC CTGGCAGGAT GTGAAGAGCG AAATGCGCAG GGCATGTGTT CGTTGGCAGG
 3061 TTTTACGCAA ATCGTGAATG AAGCACGCA ACCCGCTTGC AGTTTGTAAAG GTATAAGGCA
 3121 GTTATTGGTG CCCTTAAACG CCTGGTGTCA CGCCTGAATA AGTGTATAATA AGCGGATGAA
 3181 TGGCAGAAAT TCGCCGGATC TTGTTGAGG AACCTTACTT CTGTTGTTG ACATAATTGG
 3241 ACAAAACTTAC TACAGAGATT TAAAAAAACCT CCCACACCTC CCCCTGAACCC TGAAACATAA
 3301 AATGAATGCA ATTGTTGTG TTAACITGT TATTCAGCT TATAATGGTT ACAAAATAAG
 3361 CAATAGCATC ACAAAATTCA CAAATAAACG ATTTTTTCA CTGCAATTCTA GTTGTGGTT
 3421 GTCCAAACTC ATCAATGTAT CTTATCATGT CTGGATCGAT CCCCCGGGTAC CGAGCTCGAA
 3481 TTCGTAATCA TGGTCATAGC TGTTTCCGT GTGAAATTGT TATCCGCTCA CAATTCCACA
 3541 CAACATACGA GCGGGAAAGCA TAAAGTGTAA AGCCTGGGGT GCCTAATGAG TGAGCTAATC
 3601 CACATTAATT GCGTTGCGCT CACTGCCCG TTTCCAGTCG GGAAACCTGT CGTGCCAGCT
 3661 GCATTAATGA ATCGGCCAAC GCGCGGGGAG AGGCGGTTTG CGTATTGGC GCTCTCCGC
 3721 TTCCCTCGCTC ACTGACTCGC TGGCCTCGGT CGITCGGCTG CGGGCAGCGG TATCAGCTCA
 3781 CTCAAAGGCG GTAATACGGT TATCCACAGA ATCAGGGGAT AACGCAGGAA AGAACATGTG
 3841 AGCAAAAGGC CAGCAAAGG CCAGGAACCG TAAAAAGGCC GCGTTGCTGG CGTTTTTCCA
 3901 TAGGCTCCGC CCCCTGACG AGCATCACAA AAATGACGC TCAAGTCAGA GGTGGCGAAA
 3961 CCGGACAGGA CTATAAAAGAT ACCAGGGTT TCCCCCTGGAA AGCTCCCTCG TCGCCTCTCC
 4021 TGTTCCGACC CTGGCGCTTA CCGGATACCT GTCCGGCTTT CTCCCTTCTGG GAAGCGTGGC
 4081 GCTTTCTCAA TGTCACGCT GTAGGTATCT CAGTCGGTG TAGTCGTTG GCTCCAAGCT
 4141 GGGCTGTGTG CACGAACCCC CCGITCAGCC CGACCGCTGC GCCTTATCTCG GTAACTATCG
 4201 TCTTGAGTCC AACCCGGTAA GACAGACTT ATGCCACTG GCAGCAGCCA CTGGTAAACAG
 4261 GATTAGCAGA GCGAGGTATG TAGGCGGTGC TACAGAGTT TCAGTGTGTT GGCTTAACCTA
 4321 CGGCTACACT AGAAGGACAG TATTTGGTAT CTGCGCTCTG CTGAAGCCAG TTACCTTCCG
 4381 AAAAAGAGTT GGTAGCTCTT GATCCGGCAA ACAAAACACC GCTGGTAGCG GTGGTTTTTT
 4441 TGTTTGCAG CACGAGATT CGCGCAGAAA AAAAGGATCT CAAGAAGATC CTTTGTATCTT
 4501 TTCTACGGGG TCTGACGCTC AGTGGAAAGA AAACTCACGT TAAGGGATT TGGTCATGAG
 4561 ATTATCAAAA AGGATCTTCA CCTAGATCTT TTAAATTAA AAATGAAGTT TTAAATCAAT
 4621 CTAAAGTATA TATGAGTAAA TTGGTCTGA CAGTTACCAA TGCTTAATCA GTGAGGCACC
 4681 TATCTCAGCG ATCTGTCTAT TTGCTTCATC CATAGTTGCC TGACTCCCG TCGTGTAGAT
 4741 AACTACGATA CGGGAGGGCT TACCATCTGG CCCCAGTGCT GCAATGATAC CGCGAGACCC
 4801 ACGCTCACCG GCTCCAGATT TATCAGCAAT AAACCAAGGCC GCCGGAAGGG CCGAGCGCAG
 4861 AAGTGGTCCT GCAACTTTAT CCGCCTCCAT CGAGTCTATT AATTGTTGCC GGGAAAGCTAG
 4921 AGTAAGTAGT TCGCCAGTTA ATAGTTGCG CAACGTTGTT GCCATTGCTA CAGGACATCGT
 4981 GGTGTCACGC TCGTCGTTG GTATGGCTTC ATTCAAGCTCC GGTTCCCAAC GATCAAGGCG
 5041 AGTTACATGA TCCCCCATGT TGTGAAAAA AGCGGTTAGC TCCCTCGGTC CTCCGATCGT
 5101 TGTCAAGAGT AAGTTGGCCG CAGTGTATC ACTCATGGTT ATGGCAGCAC TGCATAATTG

Figure 20 (continued):

5161 TCTTACTGTC ATGCCATCCG TAAGATGCTT TTCTGTGACT GGTGAGTACT CAACCAAGTC
5221 ATTCTGAGAA TAGTGTATGC GGCAGCCGAG TTGCTCTTGC CGGGCGTCAA TACGGGATAA
5281 TACCGGCCA CATAGCAGAA CTTTAAAAGT GCTCATCATT GGAAAACGTT CTTCGGGCG
5341 AAAACTCTCA AGGATCTTAC CGCTGTTGAG ATCCAGTTCG ATGTAACCCA CTCGTGCACC
5401 CAACTGATCT TCAGCATCTT TTACTTCAC CAGCGTTCT GGGTGAGCAA AAACAGGAAG
5461 GCAAAATGCC GCAAAAAGG GAATAAGGGC GACACGGAAA TGTTGAATAC TCATACTCTT
5521 CCTTTTCAA TATTATTGAA GCATTTATCA GGGTTATTGT CTCATGAGCG GATACATATT
5581 TGAATGTATT TAGAAAAATA AACAAATAGG GGTTCCGCGC ACATTTCCCC GAAAAGTGCC
5641 ACCTGACGTC TAAGAAACCA TTATTATCAT GACATTAACC TATAAAAATA GGCGTATCAC
5701 GAGGCCCTT CGTCTCGCGC GTTTCGGTGA TGACGGTGAA AACCTCTGAC ACATGCAGCT
5761 CCCGGAGACG GTCACAGCTT GTCTGTAAGC GGATGCCGGG AGCAGACAAG CCCGTCAGGG
5821 CGCGTCAGCG GGTGTTGGCG GGTGTCGGGG CTGGCTTAAC TATGCGGCAT CAGAGCAGAT
5881 TGTACTGAGA GTGCACCATA TGCGGTGTGA AATACCGCAC AGATGCGTAA GGAGAAAATA
5941 CCGCATCAGG CGCCATTGCG CATTCAAGGCT GCGCAACTGT TGGGAAGGGC GATCGGTGCG
6001 GGCCTCTTCG CTATTACGCC AGCTGGCGAA AGGGGGATGT GCTGCAAGGC GATTAAGTTG
6061 GGTAACGCCA GGGTTTCCC AGTCACCGACG TTGTAAAACG ACGGCCAGTG CCAAGC

//

Figure 21: Nucleic acid sequence of the known segment of the R15/appa transgene used for the generation of transgenic mice (SEQ ID NO:5).

LOCUS R15/appa 3470 bp DNA SYN 15-APR-2000
 DEFINITION R15/appa transgene with vector sequences removed.
 ACCESSION R15/appa
 REFERENCE 1 (bases 1 to 3470)
 SOURCE synthetic construct.
 ORGANISM synthetic construct
 artificial sequence.
 KEYWORDS salivary proline-rich protein, acid glucose-1-phosphatase; appA
 gene; periplasmic phosphoanhydride phosphohydrolase; artificial
 sequence;
 AUTHORS Golovan, S., Forsberg, C.W., Phillips, J.
 JOURNAL Unpublished.

DEFINITION Rat salivary proline-rich protein (RP15) gene.
 ACCESSION M64793 M36414
 VERSION M64793.1 GI:206711
 SOURCE Rat (Sprague-Dawley) liver DNA.
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;
 Mammalia;
 Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 1748)
 AUTHORS Lin, H.H. and Ann, D.K.
 TITLE Molecular characterization of rat multigene family
 encoding proline-rich proteins
 JOURNAL Genomics 10, 102-113 (1991)
 MEDLINE 91257817
 FEATURES
 source Location/Qualifiers
 1..1748
 /organism="Rattus norvegicus"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /tissue_type="liver"
 /tissue_lib="cosmid genomic library"
 misc_feature 1802-1810
 /function=" consensus sequence for initiation in
 higher eukaryotes "

FEATURES Location/Qualifiers

DEFINITION E. coli periplasmic phosphoanhydride phosphohydrolase (appA) gene,
 ACCESSION M58708 L03370 L03371 L03372 L03373 L03374 L03375
 VERSION M58708.1 GI:145283
 SOURCE Escherichia coli DNA.
 ORGANISM Escherichia coli
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 Escherichia.
 REFERENCE 1 (bases 1811..3109)
 AUTHORS Dassa, J., Marck, C. and Boquet, P.L.
 TITLE The complete nucleotide sequence of the Escherichia coli gene appA
 reveals significant homology between pH 2.5 acid phosphatase
 and glucose-1-phosphatase

Figure 21 (continued):

JOURNAL J. Bacteriol. 172 (9), 5497-5500 (1990)
 MEDLINE 90368616

FEATURES	Location/Qualifiers			
Source	1811..3109 <i>/organism="Escherichia coli"</i> <i>/db_xref="taxon:562"</i>			
sig_peptide	1811..1876 <i>/gene="appA"</i>			
CDS	1811..3109 <i>/gene="appA"</i> <i>/standard_name="acid phosphatase/phytase"</i> <i>/transl_table=11</i> <i>/product="periplasmic phosphoanhydride phosphohydrolase"</i> <i>/protein_id="AAA72086.1"</i> <i>/db_xref="GI:145285"</i>			
<i>/translation="MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAP</i>				
TKATQLMQDVTPDAWTPVKGWLTPRGELIAYLGHYQRQLVADGLLAKKGCPQS				
GQVAIIDVDERTRKTGEAPAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNA				
NVTDAILSRAGGSIADFTGHRQTAFRELERVLFNFPQSNLCLKREKQDESCSLTQALPS				
ELKVSADNVSLTGAWSLASMLTEIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQF				
YLLQRTPEVARSRATPLLDLIKTALTPHPPQKQAYGVTLPSTVLFIAQHDTNLANLGG				
ALELNWTLPQPDNTPPGGEVLVFERWRRILSDNSQNIQVSLVQTLQQMRDKTPSLNT				
PPGEVKLTLAGCEERNAQGMCSLAGFTQIVNEARIPACSL"				
mat_peptide	1877 3106 <i>/gene="appA"</i> <i>/product="periplasmic phosphoanhydride phosphohydrolase"</i>			
mutation	replace(1817.. 1819, "gcg changed to gcc") <i>/gene="appA"</i> <i>/standard_name="A3 mutant"</i> <i>/note="created by site directed mutagenesis"</i> <i>/phenotype="silent mutation"</i>			
mutation	replace(3092..3094, "ccg changed to ccc") <i>/gene="appA"</i> <i>/standard_name=" P428 mutant"</i> <i>/note="created by site directed mutagenesis"</i> <i>/phenotype=" silent mutation "</i>			
mutation	replace(3095..3097, "gcg changed to gct") <i>/gene="appA"</i> <i>/standard_name=" A429 mutant"</i> <i>/note="created by site directed mutagenesis"</i> <i>/phenotype=" silent mutation "</i>			
 <i>polyA_signal 3262..3457</i> <i>/note="SV40 signals"</i>				
BASE COUNT	1065 a	721 c	735 g	949 t
ORIGIN				
1	GGATCCCCTT TGCTATGTAG TTTTAATGG AAATTACAAC CCATAGTGTG TTGATAAATA			
61	GAGAGTCCTG TTGGTTTAA GCAACCTCTG TTTCTATAA ACTCCATAAA AACAGGAATA			
121	CTCTTGTGTT CTAGCATAAC CAAAAGATTT AGTGAATTGA AAACAATGTT CCCTTAGAGT			
181	ATAGGTCTAA TAACCCGAA AATATTACCA TGATACTGAG CATTGTAAAG TATCTCATAG			
241	CATGTAGTAT CCATAGTCCA TCAATGAGAG AGACATTAA CATGATTTC ATTAATCAGG			

Figure 21 (continued):

301 TGGAAAAGAC ATGACAACAT TCACAGGCAC TGACAGAAC ATAGTGGTCC ACCTTGCACA
 361 TATTTCACTA AACTAGGTT ATCTATTTTG TTGCTTCTC TAACATCTCT GCAATGAAGC
 421 AGGTCAACAG TGCCACATAT CCTTTACTTA ACCTAAGGAA CACAAAAAAT TTTCTACATA
 481 TATCCTGGTT AGAGAGTGCT TAAAATAAGT TTTCCAAGAA TGGAAAAGAA ATGTTCTGAC
 541 TTAACAATTA AGACAGTATT TATTTAAAGC AAGAAATATG AGGCACACAA GAAAATATTT
 601 TGGGAAGAAA CCATTGGTG AACAAATATTT CAAATAAAA TAGACAAACA TAGTTAATTG
 661 TAAAACATAT GTTGACCAG CCCCTTTT CAATAGGCTT AATGTGAATA AAATGTTAAA
 721 GATTCTCTT GGGTGGCTGC AAATTGTCGA CGAATAAGAC AAAATATAAA AATAAGGACT
 781 GAGTCTCACA AAATGAAAAG GAAATATAATT CAGAAAGAGA ATCTTGAGAG AATGTTGTTG
 841 CACAAATTAA AGAAAACCTG TGGTGAATGA CATCTGAGG CCTGAGCTAT TACTGACATT
 901 TAAGATAAAG GTAACTGTAT ACATTTGTCC CATTGAGGGG ACAAGAAAGC TGCTCTCATG
 961 TTCAGCTCTA TAATTCTTGC CTAAACAAAC TTAAATAGAA TGATTTAAAAA TATGGAGCTG
 1021 TCCATGGACC TTGAAAATAT AAAATAGTCA AGCAACTTAT CAAGGAATTAA CAGATTCCTT
 1081 GATACTAACA CAGGTAAATC CCACACGTGT TTTGAGACTA CATTGCTGG GATTTATTG
 1141 ATGTAATAGG TCACATGTTT TTGCGGCCAA TGTTGCTGTT ATTGGTTAC TTCAAGAGAA
 1201 TAGTGGCAAC TGATGCTATG TATTCTAGGG GTTTGAAGTG ATGTTTCATG ATTGAAATTT
 1261 GTAAAAGAAT AACATCATCA TTCTTAACAA TAGAACATAT AAAGTCACAC AGAAGTGACA
 1321 GTGTTTAAGC TGACTATTG ATCAAAGAAA TTTATTACCT TCAGTTTCAA TGAAATAAT
 1381 TACTGATAAT ACAAACATGT GTGAACACAC ACTAATCCTA TCCAATGCA CAGTGATACA
 1441 CAGAAAATAT TAGCAAGTAG AATGCAATAT TTATATAACG ATTGTATTAA TCAATCAATT
 1501 GTATGTATCA ATATATGGGC TATTTCTTCA CACATGATTT TATTCAAATT TACTCTAATC
 1561 ATTGTTGAAAC CATTAGAAA AGGCATACTG GCAACTTTC CTTACCTCAT CCAGCTGGC
 1621 AAAAGTCCCAGTGGAGTA AAGGATGCAA GATTCTCTGC TCTGTTAAGT ATAAAATAAT
 1681 AGTATGAATT CAAAGGTGCC ATTCTCTGC TTCTAGTTAT AAAGCAGTG CTTGCTCTT
 1741 CCAGCACAGA TCTGGATCTC GAGGAGCTTG GCGAGATTTT CAGGACTAA GGAAGCTAAA
 1801 AGCCGCCACC ATGAAAGCCA TCTTAATCCC ATTTTATCT CTTCTGATTTC CGTTAACCCC
 1861 GCAATCTGCA TTGCTCAGA GTGAGCCGGA GCTGAAGCTG GAAAGTGTGG TGATTGTCAG
 1921 TCGTCATGGT GTGCGTGCTC CAACCAAGGC CACCGCAACTG ATGAGGATG TCACCCAGA
 1981 CGCATGGCCA ACCTGGCCGG TAAAATGGG TTGGCTGACA CCGCGCGGTG GTGAGCTAAT
 2041 CGCCTATCTC GGACATTACC AACGCCAGCG TCTGGTAGCC GACGGATTG TGCGAAGGG
 2101 GGGCTGCCCG CAGTCTGGTC AGGTGCGAT TATTGCTGAT GTGACGAGC GTACCCGTAA
 2161 AACAGGCAGA GCCTTCGCCG CCGGGCTGGC ACCTGACTGT GCAATAACCG TACATACCCA
 2221 GGCAGATACG TCCAGTCCCG ATCCGTTATT TAATCCTCTA AAAACTGGCG TTGCAACT
 2281 GGATAACCGC AACGTGACTG ACGCGATCT CAGCAGGGCA GGAGGGTCAA TTGCTGACTT
 2341 TACCGGCAT CGGAAACCGG CGTTTGCAGA ACTGGAACCG GTGCTTAATT TTCCGCAATC
 2401 AAACTTGTGC TTAAACGTG AGAAACAGGA CGAAAGCTGT TCAATTAACGC AGGCATTACC
 2461 ATCGGAACTC AAGGTGAGCG CCGACAATGT CTCATTAACC GGTGCGGTAA GCCTCGCATC
 2521 AATGCTGACG GAGATATTTC TCCCTGCAACA AGCACAGGGAA ATGCCGGAGC CGGGGTGGGG
 2581 AAGGATCACC GATTACACCC AGTGGAAACAC CTTGCTAAGT TTGCTATAACG CGCAATTTC
 2641 TTGCTACAA CGCACGCCAG AGGTGCGCC CAGCCCGCGCC ACCCCGTAT TAGATTGAT
 2701 CAAGACAGCG TTGACGCCCG ATCCACCGCA AAAACAGGGCG TATGGTGTGA CATTACCCAC
 2761 TTCAGTGTGTTIATCGCG GACACGATAC TAATCTGGCA AATCTGGCG GCGCACTGGA
 2821 GCTCAACTGG ACGCTTCCCG GTCAAGCCGA TAACACGCCG CCAGGTGGTG AACTGGTGT
 2881 TGAACGCTGG CGTCGGCTAA GCGATAACAG CCAGTGGATT CAGGTTTCCG TGGTCTTCCA
 2941 GACTTTACAG CAGATGCGTG ATAAAACGCC GCTGTCATTA AATACGCCGC CGGGAGAGGT
 3001 GAAACTGACC CTGGCAGGAT GTGAAGAGCG AAATGCGCAG GGCATGTGTT CGTTGGCAGG
 3061 TTTTACGCCAA ATCGTGAATG AAGCACGCAT ACCCGCTTGC AGTTTGTAAAG GTATAAGGCA
 3121 GTTATTGGTG CCCTTAAACG CCTGGTGTCA CGCCTGAAATA AGTGATAATA AGCGGATGAA
 3181 TGGCAGAAAT TCGCCGGATC TTTGTGAAGG AACCTTACTT CTGTTGTTG ACATAATTGG
 3241 ACAAAACTACC TACAGAGATT TAAAAAAACCT CCCACACCTC CCCCTGAACC TGAAACATAA
 3301 AATGAATGCA ATTGTTGTTG TTAACCTGTT TATTGCAAGCT TATAATGGTT ACAAAATAAG
 3361 CAATAGCCTC ACAAAATTCA CAAATAAAGC ATTGTTTCTA CTGCAATTCTA GTGTTGGTTT
 3421 GTCCAAACTC ATCAATGTAT CTTATCATGT CTGGATCGAT CCCGGGTAC

//

Figure 22: Nucleic acid sequence of the SV40/APPA+intron plasmid (SEQ ID NO:6).

LOCUS SV40/APPA 5421 bp DNA CIRCULAR SYN 14-APR-2000
 DEFINITION Ligation of SV40 promoter/enhancer into CAT/APPA+intron
 ACCESSION SV40/APPA
 REFERENCE 1 (bases 1 to 5421)
 SOURCE synthetic construct.
 ORGANISM synthetic construct
 artificial sequence.
 KEYWORDS SV40 promoter/enhancer, acid glucose-1-phosphatase; appA gene;
 periplasmic phosphoanhydride phosphohydrolase; artificial
 sequence;
 AUTHORS Golovan, S., Forsberg, C.W., Phillips, J.
 JOURNAL Unpublished.

DEFINITION E. coli periplasmic phosphoanhydride phosphohydrolase (appA) gene,

ACCESSION MS8708 L03370 L03371 L03372 L03373 L03374 L03375
 VERSION MS8708.1 GI:145283
 SOURCE Escherichia coli DNA.
 ORGANISM Escherichia coli
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 Escherichia.

REFERENCE 1 (bases 40 1337)
 AUTHORS Dassa, J., Marck, C. and Boquet, P.L.
 TITLE The complete nucleotide sequence of the Escherichia coli gene appA
 reveals significant homology between pH 2.5 acid phosphatase
 and glucose-1-phosphatase
 JOURNAL J. Bacteriol. 172 (9), 5497-5500 (1990)
 MEDLINE 90368616

FEATURES Location/Qualifiers
 Source 40 1337
 /organism="Escherichia coli"
 /db_xref="taxon:562"
 sig_peptide 40.. 105
 /gene="appA"
 CDS 40 1337
 /gene="appA"
 /standard_name="acid phosphatase/phytase"
 /transl_table=11
 /product="periplasmic phosphoanhydride phosphohydrolase"
 /protein_id="AAA72086.1"
 /db_xref="GI:145285"

 /translation="MKAILIPFLSLLIPLTPQSFAQSEPELKLESVVIVSRHGVRAP
 TKATQLMQDVTPTDAWPTWPVKGWLTPRGELIAYLGHYQRQLVADGLLAKKGCPQS
 GQVAIIDVDERTRKTGEAFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNA
 NVTDAILSRAGGSIADFTGHRQTAFRELERVLFNFPQSNLCLKREKQDESCSLTQALPS
 ELKVSADNVSLTGAVSLASMLTEIFLILQQAQGMPEPGWGRITDSHQWNNTLLSLHNAQF
 YLLQRTPEVARSRATPLLDLIKTALTYPHPPQKQAYGVTLPSTSVLFIAGHDNTNLANLGG
 ALELNWTLPGQPDNTPPGELVFERWRRLSDNSQWIQVSLVFTQTLQQMRDKTPLSLNT
 PPGEVKLTLAGCEERNAQGMCSLAGFTQIVNEARIPACSL"
 mat_peptide 106 1334
 /gene="appA"

Figure 22 (continued):

```

/product="periplasmic phosphoanhydride phosphohydrolase"

mutation      replace(46.. 48,"gcg changed to gcc")
/gene="appA"
/standard_name="A3 mutant"
/note="created by site directed mutagenesis"
/phenotype="silent mutation"
mutation      replace(1320..1322," ccg changed to ccc")
/gene="appA"
/standard_name=" P428 mutant"
/note="created by site directed mutagenesis"
/phenotype=" silent mutation "
mutation      replace(1323..1325," gcg changed to gct")
/gene="appA"
/standard_name=" A429 mutant"
/note="created by site directed mutagenesis"
/phenotype=" silent mutation "

```

DEFINITION Plasmid pBLCAT3 (bases 2200 to 4924)

ACCESSION X64409

VERSION X64409.1 GI:58163

SOURCE synthetic construct.

ORGANISM synthetic construct

artificial sequence.

REFERENCE 1 (bases 2200 to 4924)

AUTHORS Luckow, B.H.R.

TITLE Direct Submission

JOURNAL Submitted (06-FEB-1992) B.H.R. Luckow, German Cancer Res Center, Im Neuenheimer Feld 280, W-6900 Heidelberg, FRG

REFERENCE 2 (bases 2200 to 4924)

AUTHORS Luckow, B. and Schutz, G.

TITLE CAT constructions with multiple unique restriction sites

for the functional analysis of eukaryotic promoters and

regulatory elements

JOURNAL Nucleic Acids Res. 15 (13), 5490 (1987)

MEDLINE 87260024

COMMENT Promoterless CAT vector for transient transfection

experiments with eukaryotic cells. Allows the analysis of foreign promoters and enhancers.

FEATURES Location/Qualifiers

source 2200 to 4924

/organism="synthetic construct"

/db_xref="taxon:32630"

SV40 t intron 1380..1993

/note="SV40 signals"

polyA_signal 1990..2230

/note="SV40 signals"

CDS complement(3471..4317)

/codon_start=1

/transl_table=11

/gene="Amp"

/product="beta-lactamase"

/protein_id="CAA45753.1"

/db_xref="GI:58165"

Figure 22 (continued):

SV40 promoter/enhancer 5023..5402
 /note="SV40 signals"

BASE COUNT 1413 a 1321 c 1331 g 1355 t
 ORIGIN

1 CGAGATTTTC AGGAGCTAAG GAAGCTAAAA GCGGCCACCA TGAAAGCCAT CTTAATCCCA
 61 TTTTTATCTC TTCTGATTCC GTAAACCCCG CAATCTGCAT TCGCTCAGAG TGAGCCGGAG
 121 CTGAAGCTGG AAAGTGTGGT GATTGTCAGT CGTCATGGTG TGCGTGCTCC AACCAAGGCC
 181 ACGCAACTGA TGCAGGATGT CACCCAGAC GCATGGCCAA CCTGGCCGGT AAAACTGGGT
 241 TGGCTGACAC CGCGNGGTGG TGAGCTAATC GCCTATCTCG GACATTACCA ACGCCAGCGT
 301 CTGGTAGCCG ACGGATTGCT GGCGAAAAAG GGCTGCCCGC AGTCTGGTCA GGTGCGATT
 361 ATTGCTGATG TCGACGAGCG TACCCGTAAA ACAGGGAAAG CCTTCGCCGC CGGGCTGGCA
 421 CCTGACTGTG CAATAACCGT ACATACCCAG GCAGATAACGT CCAGTCCCCA TCCGTTAITT
 481 AATCCTCTAA AAACCTGGCGT TTGCCAATC GATAACCGA ACGTGACTGA CGCGATCCTC
 541 AGCAGGGCAG GAGGGTCAAT TGCTGACTTT ACCGGGCATC GGCAACGGC GTTTCGCGAA
 601 CTGGAACGGG TGCTTAATTT TCCGCAATCA AACCTGTGCG TTAAACGTGA GAAACAGGAC
 661 GAAAGCTGTT CATTAAACCA GGCATTACCA TCGGAACTCA AGGTGAGCGC CGACAATGTC
 721 TCATTAACCG GTGCCGTAAG CCTCGCATCA ATGCTGACGG AGATATTCT CCTGCAACAA
 781 GCACAGGGAA TGCCGGAGCC GGGGTGGGGA AGGATCACCG ATTACACCCA GTGGAACACC
 841 TTGCTAAGTT TGCTAAACCG GCAATTTTAT TTGCTACAAAC GCACGCCAGA GTTTCGCGC
 901 AGCCGCGCCA CCCCGTTATT AGATTTGATC AAGACAGCGT TGACGCCCA CCACCGCAAA
 961 AACAGGGCTA TGGTGTGACA TTACCCACTT CAGTGTGTT TATCGCCCGA CACGATACTA
 1021 ATCTGGCAAA TCTCGGGCGC GCACCTGGAGC TCAACTGGAC GCTTCCCGT CAGCGGATA
 1081 ACACGGCGCC AGGTGGTGAA CTGGTGTGTT AAGCCTGGCG TCGGCTAACG GATAACAGCC
 1141 AGTGGATTCA GGTTCGCTG GTCTTCCAGA CTTTACAGCA GATGCGTGAT AAAACGCCGC
 1201 TGTCTTAAAC TACGCCGCC GGAGAGGTGA AACTGACCCCT GGCAAGGATGT GAAGAGCGAA
 1261 ATGCGCAGGG CATGTGTTCG TTGGCAGGTT TAACGCAAAT CGTGAATGAA GCACGCATAC
 1321 CCGCTTGCAG TTGTAAGGC AGTTATTGGT GCCCTTAAAC GCCTGGTGT ACGCCTGAAT
 1381 AAGTGATAAT AAGCGGATGA ATGGCAGAAA TTGCCGGAT CTTTGTGAAG GAACCTTACT
 1441 TCTGTGGTGT GACATAATTG GACAAACTAC CTACAGAGAT TIAAAGCTCT AAGGTTAAATA
 1501 TAAATTTTTT AAGTGTATAA TGTGTTAAC TACTGATTCT AATTGTTGT GTATTTTAGA
 1561 TTCCAACCTA TGGAACTGAT GAATGGGAGC AGTGGTGGAA TGCTTAAAT GAGGAAAACC
 1621 TGTTTGCTC AGAAGAAATG CCATCTAGT ATGATGAGGC TACTGCTGAC TCTCAACATT
 1681 CTACTCCTCC AAAAAGAAG AGAAAGGTAG AAGACCCAA GGACTTTCTC TCAGAATTGC
 1741 TAAGTTTTT GAGTCATGCT GTGTTAGTA ATAGAACTCT TGCTTGCCTT GCTATTACCA
 1801 CCACAAAGGA AAAAGCTGCA CTGCTATACA AGAAAATTAT GGAAAATAT TCTGTAACCT
 1861 TTATAAGTAG GCATAACAGT TATAATCATA ACATACTGTT TTTTCTTACT CCACACAGGC
 1921 ATAGAGTGTGTC TGCTTAAAT AACTATGCTC AAAAATTGTG TACCTTTAGC TTTTTAAATT
 1981 GTAAAGGGGT TAATAAGGAA TATTGATGT ATAGTGCCTT GACTAGAGAT CATAATCAGC
 2041 CATAACCACAT TTGTTAGAGGT TTTACTTGCT TIAAAAACC TCCCACACCT CCCCTGAAAC
 2101 CTGAAACATA AAATGAATGC AATTGTTGTT GTAAACTGTT TTATTGCAAG TTATAATGGT
 2161 TACAATAAA GCAATAGCAT CACAAATTTC ACAAAATAAG CATTTTTTTC ACTGCATTCT
 2221 AGTTGTGGTT TGCTTAAACT CATCAATGTA TCTTATCATG TCTGGATCGA TCCCCGGGTA
 2281 CCGAGCTCGA ATTGTAATC ATGGTCATAG CTGTTCTCG TGTAATTG TTATCCGCTC
 2341 ACAATTCCAC ACAACATACCG AGCCGGAAGC ATAAAGTGT AAGCCTGGGG TGCTTAATGA
 2401 GTGAGCTAAC TCACATTAAT TGCGTTGCGC TCACCTGGCCG CTTTCCAGTC GGGAAACCTG
 2461 TCGTGGCCAGC TGCTTAAATG AATCGGCCAA CGCGCGGGGA GAGGCGGGTT GCGTATTGGG
 2521 CGCTCTTCCG CTTCTCGCT CACTGACTCG CTGCGCTCG TGTTTGCCT GCGGCGAGCG
 2581 GTATCAGCTC ACTCAAAGGC GGTAAATACGG TTATCCACAG AATCAGGGGA TAACGCAGGA
 2641 AAGAACATGT GAGCAAAGG CCAGCAAAGG GCCAGGAACC GTAAAAAGGC CGCGTTGCTG
 2701 GCGTTTTCTC ATAGGCTCCG CCCCCCTGAC GAGCATCACA AAAATCGACG CTCAAGTCAG
 2761 AGGTGGCGAA ACCCGACAGG ACTATAAAGA TACCGGGGT TICCCCCCTGG AAGCTCCCTC
 2821 GTGCGCTCTC CTGTTCCGAC CCTGCCGCTT ACCGGATAAC TGTCCGCCCT TCTCCCTTCG
 2881 GGAAGCGTGG CGCTTCTCA ATGCTCACGC TGAGGTATC TCAGTTCGGT GTAGGTGCGTT
 2941 CGCTCCAAGC TGGGCTGTT GCACGAACCC CCCGTTCAAGC CCGACCGCTG CGCCTTATCC
 3001 GGTAACATAC GTCTTGAAGTC CAACCCGGTA AGACACGACT TATGCCACT GGCAGCAGCC
 3061 ACTGGTAACA GGATTAGCAAG AGCGAGGTAT GTAGGCCGTG CTACAGAGTT CTGGAAGTGG

Figure 22 (continued):

3121 TGGCCTAACT ACGGCTACAC TAGAAGGACA GTATTTGGTA TCTGCGCTCT GCTGAAGCCA
 3181 GTTACCTTCG GAAAAAGAGT TGGTAGCTCT TGATCCGGCA AACAAACAC CGCTGGTAGC
 3241 GGTGGTTTTT TTGTTTGCCTAAGCAGATTACGCGCAGAA AAAAAGGATC TCAAGAAGAT
 3301 CCTTGATCT TTTCTACGGG GTCTGACGCT CAGTGGAACG AAAACTCACG TTAAGGGATT
 3361 TTGGTCATGA GATTATCAAA AAGGATCTTC ACCTAGATCC TTTTAAATTA AAAATGAAGT
 3421 TTTAAATCAA TCTAAAGTAT ATATGAGTAA ACTTGGTCTG ACAGTTACCA ATGCTTAATC
 3481 AGTGAGGCAC CTATCTCAGC GATCTGTCTA TTTCGTTCAT CCATAGTTGC CTGACTCCCC
 3541 GTCGTGAGA TAATCTACGAT ACGGGGAGGGC TTACCATCTG GCCCCAGTGC TGCAATGATA
 3601 CCGCGAGACC CACGCTCACC GGCTCCAGAT TTATCAGCAA TAAACCAGCC AGCCGGAAGG
 3661 GCCGAGCGCA GAAGTGGTCC TGCAACTTTA TCCGCCTCCA TCCAGTCTAT TAATGTTGC
 3721 CGGGAAAGCTA GAGTAAGTAG TTTCGCGAGTT AATAGTTGC GCAACGTTGT TGCCATTGCT
 3781 ACAGGCATCG TGGTGTACCG CTCGTCGTTT GGTATGGCTT CATTCTAGTC CGGTTCCCAA
 3841 CGATCAAGGC GAGTTACATG ATCCCCCATG TTGTGCAAAA AAGCGGTTAG CTCCCTCGGT
 3901 CCTCCGATCG TTGTCAGAAG TAAGTGGCC GCAGTGTAT CACTCATGGT TATGGCAGCA
 3961 CTGCATAATT CTCTTACTGTCATGCCATCC GTAAGATGCT TTTCGTTGAC TGGTGTAGTAC
 4021 TCAACCAAGT CATTCTGAGA ATAGTGTATG CGGGCGACCGA GTTGTCTTG CCCGGCGTCA
 4081 ATACGGGATA ATACCGCGCC ACATAGCAGA ACTTTAAAG TGCTCATCAT TGGAAAACGT
 4141 TCTTCGGGGC GAAAACCTTC AAGGATCTT CCGCTGTGAT GATCCAGTTC GATGTAACCC
 4201 ACTCGTGCAC CCAACTGATC TTTCAGCATCT TTTCAGCATTC CCAGCGTTTC TGGGTGAGCA
 4261 AAAACACAGGA GGCAAAATGC CGCAAAAAAG GGAATAAGGG CGACACGGAA ATGTTGAATA
 4321 CTCATACTCT TCCCTTTCTA ATATTATTGA AGCATTIATC AGGGTTATTC TCTCATGAGC
 4381 GGATACATAT TTGAATGTAT TTAGAAAAAT AAACAAATAG GGGTTCCGCG CACATTCCC
 4441 CGAAAAGTGC CACCTGACGT CTAAGAAAACC ATTATTATCA TGACATTAAC CTATAAAAT
 4501 AGGCATGATCA CGAGGCCCTT TCGTCTCGCG CGTTTCCGGTG ATGACGGTGA AAACCTCTGA
 4561 CACATGCAGC TCCCAGGAGAC GGTACAGCT TGTCTGTAAAG CGGATGCCGG GAGCAGACAA
 4621 GCCCCGTCAAGG GCGCGTCAGC GGGTGTGGC GGGTGTGGG GCTGGTTAA CTATGCGGCA
 4681 TCAGAGCAGA TTGTAATGAG AGTGCACCAT ATGCGGTGTG AAATACCGCA CAGATGCGTA
 4741 AGGAGAAAAT ACCGCATCAG GCGCCATTG CCAATTCTAGGC TGCGCAACTG TTGGGAAGGG
 4801 CGATCGGTGC GGGCCTCTTC GCTATTACGC CAGCTGGCGA AAGGGGGATG TGCTGCAAGG
 4861 CGATTAAGTT GGGTAACGCC AGGGTTTCC CAGTCACGAC GTTGTAAAAC GACGCCAGT
 4921 GCCAAGCTTT ACACCTTATG CTTCCGGCTC GTATGTTGTG TGGATTGTG AGCGGATAAC
 4981 AATTTACAC AGGAAACAGC TATGACCATG ATTACGAATT CGGCGCAGCA CCATGGCCTG
 5041 AAATAACCTC TGAAAGAGGA ACTTGGTTAG GTACCTTCTG AGGCGGAAAG AACCAAGCTGT
 5101 GGAATGTGTG TCAGTTAGGG TGTGGAAAGT CCCCAGGCTC CCCAGCAGGC AGAAGTATGC
 5161 AAAGCATGCA TCTCAATTAG TCAGCAACCA GGTGTGGAAA GTCCCCAGGC TCCCCAGCAG
 5221 GCAGAAGTAT GCAAAGCATG CATCTCAATT AGTCAGCAAC CATAGTCCCG CCCCTAACTC
 5281 CGCCCATCCC GCCCCTAACT CCGCCCCAGTT CCGCCCCATTC TCCGCCCCAT GGCTGACTAA
 5341 TTTTTTTTAT TTATGCAGAG GCCGAGGCCG CCTCGGGCTC TGAGCTATTG CAGAAGTAGT
 5401 GAGGAGGCTC GAGGAGCTTG G

//

Figure 23. The nucleic acid sequence of the Lama2/APPA transgene used for the generation of transgenic mice and transgenic pigs (SEQ ID NO: 7)

LOCUS transgene 17732 bp DNA SYN 14-APR-2000
 DEFINITION Lama-appA cut XhoI..20623 to NotI..17732
 ACCESSION transgene
 KEYWORDS parotid secretory protein; acid glucose-1-phosphatase; appA gene; periplasmic phosphoanhydride phosphohydrolase; artificial sequence; cloning vector
 REFERENCE 1 (bases 1 to 17732)
 AUTHORS Golovan, S., Forsberg, C.W., Phillips, J.
 JOURNAL Unpublished.

FEATURES

DEFINITION M. musculus Psp gene for parotid secretory protein.
 ACCESSION X68699
 VERSION X68699.1 GL:53809
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 3777 to 5332;)
 AUTHORS Svendsen, P., Laursen, J., Krogh-Pedersen, H. and Hjorth, J.P.
 TITLE Novel salivary gland specific binding elements located in the PSP proximal enhancer core
 JOURNAL Nucleic Acids Res. 26 (11), 2761-2770 (1998)
 MEDLINE 98256451
 REFERENCE 2 (bases 7147 to 12653; 13952 to 17731)
 AUTHORS Mikkelsen, T.R.
 TITLE Direct Submission
 JOURNAL Submitted (07-OCT-1992) T.R. Mikkelsen, Department of Molecular Biology, University of Aarhus, CP Mollers Alle 130, 8000 Aarhus, DENMARK
 REFERENCE 3 (bases 7147 to 12653; 13952 to 17731)
 AUTHORS Laursen J, Hjorth JP
 TITLE A cassette for high-level expression in the mouse salivary glands.
 JOURNAL Gene 1997 Oct 1;198(1-2):367-72
 MEDLINE 9370303

FEATURES

Location/Qualifiers

source 1..to 12653; 13952 to 17731
 /organism="Mus musculus"
 /strain="C3H/As"
 /db_xref="taxon:10090"
 /chromosome="2"
 /map="Estimate: 69 cM from centromere"
 /clone="Lambda YP1, Lambda YP3, Lambda YP7"
 /clone_lib="Lambda-PHAGE (Lambda L47.1)"
 /germline
 /note="Allele: b"
 misc_feature 3777-5332
 /gene="PSP"
 /function="salivary gland specific positive acting regulatory region"
 enhancer 7147..8724

Figure 23 (continued):

```

  /evidence=experimental
exon 11778..11824
  /gene="Psp"
  /note="exon a"
  /number=1
  /evidence=experimental
exon 12626.. 14190
  /gene="Psp"
  /note="exon b fused with exons h and i"
misc_feature 12644-12652
  /function=" consensus sequence for initiation in higher
  eukaryotes "
misc_feature 13952-13965
  /function=" M13mp18 polylinker"

```

DEFINITION *E. coli* periplasmic phosphoanhydride phosphohydrolase (appA) gene,

```

ACCESSION M58708 L03370 L03371 L03372 L03373 L03374 L03375
VERSION M58708.1 GI:145283
SOURCE Escherichia coli DNA.
ORGANISM Escherichia coli
Bacteria; Proteobacteria; gamma subdivision;
Enterobacteriaceae;
Escherichia.

```

```

REFERENCE 1 (bases 12653..13951)
AUTHORS Dassa,J., Marck,C. and Boquet,P.L.
TITLE The complete nucleotide sequence of the Escherichia coli
      gene appA reveals significant homology between pH 2.5
      acid phosphatase and glucose-1-phosphatase
JOURNAL J. Bacteriol. 172 (9), 5497-5500 (1990)
MEDLINE 90368616

```

FEATURES	Location/Qualifiers
Source	12653..13951 /organism="Escherichia coli" /db_xref="taxon:562"
sig_peptide	12653..12718
/gene="appA"	
CDS	12653 13951 /gene="appA" /standard_name="acid phosphatase/phytase" /transl_table=11 /product="periplasmic phosphoanhydride phosphohydrolase" /protein_id="AAA72086.1" /db_xref="GI:145285"

```

/translation="MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAP
TKATQLMQDVTPDAWPTWPVKLGWLTPRGGELIAYLGHYQRQRQLVADGLLAKKGCPQS
GQVAIIDVDERTRKTGEAFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNA
NVTDAILSRAGGSIADFTGHRQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPS

```

Figure 23 (continued):

ELKVSADNVSLTGAVSIALSMLTEIFLLQQAQGMPEPGWGRITDSHQWNNTLLSLHNAQF
 YLLQRTPEVARSRATPLLDLIKTALTPHPPQKQAYGVTLPSTVLFIA GHDTNLANLGG
 ALELNWTLPQPDNTPPGELVFERWRRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNT
 PPGEVKLTLAGCEERNAQGMCSLAGFTQIVNEARIPACSL"
 mat_peptide 12719 13948
 /gene="appA"
 /product="periplasmic phosphoanhydride
 phosphohydrolase"
 mutation replace(12659.. 12661,"gcg changed to gcc")
 /gene="appA"
 /standard_name="A3 mutant"
 /note="created by site directed mutagenesis"
 /citation=[3]
 /phenotype="silent mutation"
 mutation replace(13934..13936, " ccg changed to ccc")
 /gene="appA"
 /standard_name=" P428 mutant"
 /note="created by site directed mutagenesis"
 /citation=[3]
 /phenotype=" silent mutation "
 mutation replace(13937..13939, " gcg changed to gct")
 /gene="appA"
 /standard_name=" A429 mutant"
 /note="created by site directed mutagenesis"
 /citation=[3]
 /phenotype=" silent mutation "

BASE COUNT 4719 a 4125 c 4168 g 4719 t
 ORIGIN

1 TCGAGAGTAT CTTTGTCAAG TGTGCCTCCA ACAAAAGGGGT ACTGTTGCC ACATAGAAAG
 61 ATCTAAACTA ATTAATTAAT CCCTCACCCG CAAATCTTTC AGTCACTAAG TTAGCACGAT
 121 TGTGAAACAA GTTCTCCAAA GGAGAGATAC AGATGAGTGC GTATAGGGTG GACCTGGCTG
 181 CTGAGGAGAC ACCTGCATCT GACTAAGAAG AGCCACGGTG TTAGTTGAAT GGTGTGGAGT
 241 AGGGTGGTTC TGTGGGACAG TAGAAAATCG AGAGGCATGT GCCGTTTAGT GAACTGATGG
 301 AAGCTACCCC AAACGACAGA GATTGTCAGT CAGGCCAATC CGTTTCGAGT TTGATGGGCA
 361 GCGGGACAGT GAGACAGACAC CACCTACTCA GTTGGAGGAA GGATGAGAAC AATGCCAGC
 421 AGGGATTGAG AGACCCCTGAC AGGCGCAAGG CCCTAACACA CACACCTTAC ACCTTACITG
 481 ACAAAAGCTGC CAAAGACCAA AGACTTGTTC TCCATTAGAA ATGACAGCTG GCTTGGACCCG
 541 ACAGCATAAT AAGCAGAGTG TACTCTGATT GGAGAACTTT AATGTGTITC ATTCACTTATT
 601 ATAAAAGGAC AGTATTACAG ATTTTGTGT ACACTGCTGT TACATGTGGG GCAGTGTGTC
 661 TTTAAAGTAGG GTAAAGTACT CTTTAAAAT GGGTCCCTAGA TATTTTTTCC TTTAACTCAA
 721 GTCTCTTACT GTTTAAATGA TTTTTATTIT GTTTAATATG GAGGAAAAAG AAGCGTAAAT
 781 GGACAATATA TATTTAGAGA AAGATGTTA GCTGTCAGAA AAATATGCAA ATCAAAATCA
 841 CACCAAGACT GCAGCACACC CCTGTCAGAT GGCTGTGATC AAGAAAATAA ATGACAATGA
 901 GTGGTGGTGA AGATGTACTA AAGGGAAACA CACACACACA CACACACACA CACACACACA
 961 CACACTGGAG CAACCACTGT GGAAATCAGT ATGAATGGTC CTCAAAACC TGAAGATAGA
 1021 GCGGGGCGTG GTGGCATACA CTTTTATITCC CAGCACTGGG GAGGCAGAGG CAGGTGGATC
 1081 TCTGAGTTCC AGGCCAGCCT GGTCTATAGC ACAGGTTCTA GGACAGCCAG GGCTACACAG
 1141 AAAAACCTG CCTTGATTAA ACCAAACCAA ACCAAACCAA ACCAAACCAA ACCAAACCAA
 1201 ACCAAACCAA ACCAAACCAAG ACCAAACCAA AACACTGAAG ATAGAACTTC AGTATTCCAT
 1261 TCCTAGATAT ATACCCAATG GAGACTAAGT CAGCAAGACAC CCTGCACAGC CATGTTCACT
 1321 ACTACACTGT TCACCCACAGC CAGGCTGTGG AACCAACCTG AGTGTCCATG ATAAATGAAT

Figure 23 (continued):

1381 GGATAGGTAA CTTTCAAGGT AAATGGACTC TGCTGTGTAC ATGCCTCACA TTCTGTTAT
 1441 TCATTTTCTT TTATGAGGTG TCCATTCAAGG AGTCACATGG TAGTTCTATT TTCAAGTCTTC
 1501 TGAAGATACT ACACTGGTCC CCACAGTTA CACTTTATC AGCAGTGAAT AAGGGTCCCT
 1561 CTATCCTTAC CATCATTGTG TGTAATTCTT CTTGATGACC CTCTTTCTGA CAGGGATAGG
 1621 ATGTAATATC AGTGTGAGGA AGTACAACCT GTTTTCTAAG TATTTATTGG CCCCTTGCAT
 1681 TTCTTCTT GAAAAGTGC GGTTTCTGAC ATCTGCTCAG GTATTCTATG GATGTGTTT
 1741 CTTGGGTGTT TGAGTTCTTA TGAATTCTAG ATGTTAAATC CCTGCCTGTG TTCTCTCCC
 1801 ATTCTGTAGG CTGCCTCCCT ACCCTGGCAA TTGTTGTCCCT TGTTTGTGAG AAACCTTTGA
 1861 TTTCATGGAA TCTCATTGTG CAGTTTCCC TCCTCTGCTA TAGCCTGAGC TAATGCACTG
 1921 GTTTTACAG AGCCCTGGTC TATGCCTTTA TCCTCCCTTG GCAGCTTCGG AGTTTCAATT
 1981 CTTACATTTA GATCTTTGAT CCACCTTGAA CAAGTTTGG AGCAGGGTGA GAGATAACGAA
 2041 TCTAGTTCCA TTCTTCCATA TGTGATCCTA GTTTACATAG CATCGTTGGT TGAAGAGGTT
 2101 TTATTTTATT TTAAATAAT GTGTCATAAA AAACGAGGTG GTTGTAGCAG TGTGGATTG
 2161 TTCTTGTGTC CTTGATCTA CAGGTCTTGT TTTGTGTCAG TCTCATGATG TTATTTATGCT
 2221 ATGGCTCTGT CATAACAGTCT GAGGTCAAGGT ATTGTGATAT ACCTTCAGTA TTGCTCCCTC
 2281 AGACTCAGGT TTGCTTGGC CAGGAGTCAT CTTACTCAGT GCTCTTAGAG CTCCCCCAGC
 2341 ATGTAAGCTGC TACTATTCTT AGTTGATAAA TCAGGAAACT GGGGCTCAGA GAGATTAAC
 2401 GTCTTGAAC ACTTCTGGGG AGGTGAAACG TGGAGACACT AAACGTGTT TACCCGTAC
 2461 TGCTCCAGTA GCTGTCGGGT GCTGGGCTAC AGCAAAGCAC CTATACTATA TATTACTCAG
 2521 GAGGTGGAAA AACTCAGCCT CCCTTGGGT TCCCAAGCTC CCAGGTGTCC AGTCACTGCT
 2581 GGAAACCTCA TGGAGTCTGA AAGGAAGGGT TGAGGGTACA TGGGGCAGGC ATGAGGAGCC
 2641 TGGGGCTGGG ATCTCCAAA CACCTGGATA TCCAGATGCC ACTGGGTCAAG GGGGAGTTGG
 2701 GAACAGAGTT GGGATGTCCA TGGACCTGTG ACAAGGCCAG GGCCAGGGGG AGGATAACTC
 2761 TGGCTTTACT AATTTGCGAA AGTCCTTAGC TTAGCAGCAG TTGTCCTGGGA GCACAGAGGG
 2821 GCCTTCTGTA AGAGGCTCAG GCAGTGC CGC TCTGTAGGCG AAGGTCTTCT CCATGTTCCC
 2881 CATGGTGGTT CTTGATGAAA GAGACAGTCC TTGGCTCAA ACTGGTTTAT TGATTGTTCA
 2941 TTGTGGAAA TGGGTGCACA CCACCTCTC AGGGTGGACAG AGAGATCAA TACCTTTGCT
 3001 AGGGAGGAAT ATCTGGGAAG GGACGCTTAC TGGCTAAACC CTCAGGGCT CTAGATACAT
 3061 CATTAGCATG GAGAACTCTG TTCTGGCTA CATGACCACA GGCCACATT CCACAAGCCA
 3121 CATGTGGAA GTGTGGCACA TGTTCAGGC CAGGAATCTG GTAGGGAGCG TGGAGCCACC
 3181 TACCACTCCA GGTGGGTGCC TGGGTGCCAG GGACCTGAA CCCGCTCAAC CTTACCAAGT
 3241 TTCCCTGGCAG GGTCCACTGT CCTACACAGA AGCTGGAGGA GGTGTGAGGG TTGTCCTT
 3301 GTGGAATGTC CCATGCTGCT TGGGGCTCAG TTCTCCACC TGTACCTCAT TGGTTGGGT
 3361 ATAAAAAGTG GGGATACTTT ATTATTCTCT GACTGGTCC TGAGGAAAAA GCATCGTGGC
 3421 AGTCCAGGAA CCACACCTG AGGTTCTGC ACTGAAGGGG CTCCTTAAGT CTCTGGAGTC
 3481 TCTCCCCCTC ACAGAGCTGC CAAAGTCTAG GTTCTTTGA GGATAACAGA GCCATGCTTG
 3541 GTAAGCAGAC AACAGCATT GTTTACTCAA CCTTCTTTG TCAAGCTCCCT TTTCATAAAC
 3601 AAGTTGAGAC ACCATGCTGG TTGAGGAAG ACTTCTAAAG CCAGACAAC GTGCAAGGAA
 3661 GAAGAAGAAG GGGCAAGTGG AGTTAGCTG GATGTAGCCC TCAAAGTCTC CAGAGACCAG
 3721 CCATGAAGGC TCAAGTGGAG GGCAAGACCT GCAGCAGCCA AGCATCTGGC AGGAGAGGAT
 3781 CCTGGGAACCC CCTCTACCAT GACACACATT TTCTCTGCGAG GTCACACTTA ATAGGCCATT
 3841 TCTTATTGTC ATCTATCATG GTGTTCTG CGAGATTAAT GAGGTGTAT GCTGCGAAC
 3901 GAAAGTATA TAAAAACAAG TCCCCCCCCC TTGTCACTGC TGCTAAGAAT GTAGCAGAAA
 3961 TTGTCCTCAAG TGTCTCTA ATCAGAAACA ATAAAGGTCT CCTTGGATTC AAGCCCTCCA
 4021 GTTCTCTCTC TCCCTGCTGA GCCTTGGACA CCCATACAAA CCTCCCTGGAT GCTACAGCTC
 4081 TGGGAGAGA CTCCAAGGTG GGGAGAGACT GATGGTACAA AAGCAAATA TTGTTGGG
 4141 GGTACACCCA CTCCCTGCTC TGTGTGGTTC CTGAGTCAG TCCCTGCAGAC AGGCCCTCAG
 4201 TGGGTCTTCC ATGGGAACA CGCAGAGGGG GGCAATGGAT GGGAAATACCC ACACCCCTGGT
 4261 TAGTTTACCC CGGCCATGCT CTCTGCTCTT CATCCCTCCT CTGCCCCCTG CCACGGCTTT
 4321 CTCTGCGAGGA ATCATATCTT CATATTGGCC CACAGGTGTT CTCTCACCC TAGCTATGAT
 4381 GTTTACTTTA GAGTGAACCTT AGCAGGGCTG GTGGGAATGA GTTCTAGAAG GCTCACGGAG
 4441 ATGCTAGGA AGAAACGTCT TCTAACTACT GAGGTACTA AGTTCTGTT GTGTCCTCT
 4501 GCCTTCCCT TGTAAAGTC ACCTTGAAAGT TAGTGCAGAA GAAATCAGAG CCCAGTCACA
 4561 GAGTAATAT GGTCTGAAAG ATTTCTTGT AGTGCAGGAA ATCCATGACA TTTCAGAGGC
 4621 CCTCTTGTGTA CCTTAAAGTC TTTGGGGTTG TATTTCTGC TTGATGTAAG TGTGTGTT
 4681 TATCAAAGAG TGAGATGGTT ACATAAGAGG TGCTCTAAAG GACAGAGAGG ATTTCAGAATT
 4741 GTGGCATGTG ACATCCTCAG GCCTTGTCT GTGGCCAGGA GGAACGTGATG CAGAAAAGAG
 4801 TAAGAGGTCA TTTCCTGGAG GCTGTCAGTA TAGAGGAGAT CTTACAGTGC ATTCCCTCCT

Figure 23 (continued):

4861 CCAGGCCCTG CCTGAGGATA GACATGTGCT GACTGCACT GAAACAGAGG CTTGGGATGG
 4921 AGAGTTAGGT TCACAGAAGG GAGGGTGGGA GATGGATGCT TGCTGGGTTG TGGGTCTCAT
 4981 CACCAGCTCC TGACCACCCG GTCAGCCAT GTGCTTATTG CATAGCTTTC TTTGCTATG
 5041 TTTACTCAGT GTGGTGTITG TTGGGACCCA GCAGAAGCCA GTCCCAGGCT GACAGCTGTG
 5101 GATACACAGG GCAGCATGAG GGTCTCAGC CTGAAGCAGT CAGGCTGGCA GAAGAGAAAG
 5161 ACCAGCACAC ATTCCCTCAA CCAACTATGT TTGAAAAAC AAACATATTA TATCACATAT
 5221 ATTGCATTAA TGAGACAGCT AAAATGACT CGGGTAGCAT GACTCCAGGT GGGGATATCT
 5281 GCAAGTGCCA TGAGTGGCAAG AGGGACAGCC AATGTGAGGC AAGAAGGAAT TCTGGCTCAA
 5341 CACAGCTTAG CTCCCTGGTG TTGGTCAAA CTTGAGAGT TTGACCACAA GCACTTTATT
 5401 TTTGACATAT TAAACAGAG CACAACCTTG GGAAAAAGTT TTCTTATGAA ATTATCACA
 5461 ATAAAGCTTA AGGCATGACT ACATTAATGC GCTCTTGCAA AGTATATGTG CCTCTTCCA
 5521 CAAGAATGGT TCTATTGACT GAGAAATAAT GTTCAGGATA AAGATCCAGG AAGAAAAGAT
 5581 CAGGGATAAG TAAATACTA AACTCTTTC CAAAGTACAT AGACCCTTT TCATAACAAT
 5641 GGGTTCTATT GACTGACAAG CACTGCTCAG GAGTTGGAA AGAGTCTAGC ATAAGCACGA
 5701 TAGCCTGGAG ACTCTAGTGA GGTCTAGTCT TACAGACAGC AAAAATCACC AGGTTACAAA
 5761 CTACATTCTCAT TTCCAGTTT CTGATCAGGC ACAGGTATGAA ATCCCTCTG TTGAAGAGAA
 5821 AAGTCCATGT TTAAATATAA TCTGGTTTCT CCAGTGTCTAT TAGCGAGAAG ACTTGAGCCC
 5881 TATACAACTC CCACCTGGAG TGACATCTG TCTCATGGT ATATTACATA CCTAGACACG
 5941 CTCATCTCAC AGACTTAGGA CTTTGCTCTG TGATCTCCAT TTCTGATCCC ACTTCCACCT
 6001 TTGCCTTGAT AGTGTCAATT TCTTCACTGC CTGGTGACA ACCATGTAT CCTCTGTGTA
 6061 TTTGAGTGT ACCATTTC GATTTTACCT GTATGCAAGA TCACACAGTC TTTGCTTTTC
 6121 TGTCTGGATG CATGCTAAC TCTACACAA AACCTTCCC CGTCACTCAG ATCTTCCCTCC
 6181 ATTAACACAT ACATGGTGT GAAAGAGGCTA GGGAGCTTCC TTTCAGTGGG GAGCTAGCTG
 6241 GCTATTGGC CTTTTTGACT GTCCAGGAAG GCCCCCAATT GCTGAGACAA GAACCTAGAT
 6301 TCTTCATTAT TGACTCTAAC TCATGTATCA AGCAGAAGCT AATGAATAGT TATCAACAGG
 6361 ATCAGAGGTT CCAGTGTAAAG ACACCTTGAC ATGAAAGAAC GGAGGAAGGA CAGATGGATG
 6421 CATAAAAGCA GGACCACTGC CCCAGGAAGG TCTGGAAAC TGATGCAGGG CAAAGGACAG
 6481 GTTATAAACCC AAATCTTAGG GAGTCAGGAAG GAGCACAGAG GAGCTAACCC AACTGACCAC
 6541 TGCTTAGGGG CTACCAACCC AATCTCCCT GTGGGAACAG CTAAGCTATC AGCCAAGGGT
 6601 AATAAACAGG CAGGACCTGT GGATGACATG GAGGACATAG GGACCTGGG TCCAGCTTT
 6661 AGCACCTGCA CTCTCAGGAT ACTCCACCAT TGTGTCTTAG AGAGCCTAGG GATACTGGGT
 6721 CCAGCCTTGC GTACCTCTCAC TCTCAGGGTA CCCCATCACT GTGTCTTGGG GAGCCTAGGC
 6781 ACCCTGGTGC CAGCCTTCAG TACCTGCGCT CTCAGGACAC CCCACCAATT TCTCTGCC
 6841 CGTCTCTCTC TTCTCTTCTC CCCCTTCATT GTCTCTCTC TTGTTCTTTC TTGACTCTCC
 6901 TTCCCCCTCA CACOCTCACT CTAGTTCTCC CCTTCCCCCTCT CTGCATCACCC TATTTCTCTC
 6961 TGTGGTCCCC CCTACTTCTC TTATCTCTCA TGCTCTCTC CTCCCCCAAA TACTTGTAC
 7021 CCACTATACT TCAGGGGCCA GCTCTAGTGA CAAAGCTGTT AATAGCAAGA CTCTCAGATC
 7081 TCCAACGGCT CAGAGGAGCC AGACCCACCA AGAAACTCTCT CCAGGTCCAA TTTCAGGTT
 7141 TTTCGAAAGC TTTCAGCAAA TGCTCAGGGAA ACATGCCACT AACAAAGAAGA TGCAAATTCC
 7201 AGTTGAGAGT GGGAAAGGCC TTGGCGTAGG TCCCATCTC CAGGCCAAGG TCAGAGGGGC
 7261 TCTGTGTAAT CCGGATTGAC AGGGCTCAGA ACAATGTTT GTTTTTAAGG TTATTTATT
 7321 TTAGGTGTTA GTGTCTTCTG TTGCTATGACC TTATGTGCAAT CATGTGTGTC CAGGTTCTG
 7381 ATGACAGTAG AGGAGGGCTT TGAATCCCTG GGGATAGGAA GTTACAGGAA ATTATAAGCT
 7441 GCTTTGTGGG TCTTCTAGCT TTCCCAACAG AAGTGAATGC TCTTCACCAC TGAGCCATCT
 7501 CTCTAGGCCA AAGAGACATT GCTTATGGA TATAATTGTC TGTGTGTGTC AACATTGAGG
 7561 AAAGGAAAT AAAAAAAA CTTCAGCCGC TAAGTTGTA CAGTTCACT AATGCTACT
 7621 TTAGTTGTG ATAAAATGGC AGGTGCTTCA ACATTTATAT ATACAAAAAC TTCCCTGCTG
 7681 GTGGTCAAC TGTGAGAACT GGGGTAAGTG GGTGAGTTCT CTTTTCTGT CTCTGCTCT
 7741 GTCTCTCTCC TTCCATTCTC TTCTAAAGGA AATAAACATT GCAGCTGGGT TATAGCTCAT
 7801 CAATATGGAA GTTACAGAAG TGAAAAAAGG CATTGCCCTG GTGGGTGGTG TTACCACTG
 7861 ATTTTTGGTT GTCCCTGCAAG GAGGTCTGGG GACTGGCTGC TCTGTCTCTG TCTGTATGAG
 7921 TGAGGGAAAGT CTGGGGAGCA GATTCCCTAA CCTTCAGGCTT GGCCTGGTTC CTGAGTGAAC
 7981 CCAGCCTCTC TGGTCTAGT AGCTTTTCC AAACAGGAAT CTGAGTGGTG ACAGGGAAACA
 8041 AGTACCAAGCC CATTGCTTAA GTGCCAGGGT TAGTGAGGGC AGGAAGCTGC CATAGCTGGG
 8101 ATTAGTAGTT GTATTGGATG TAGGAAGTCC TATCCTGGGA CAGCTAATCC TTAATGCTTC
 8161 ACTGGAGATT TTCAATGAGA AATTATCCC ACGGCCCATC TGGCCCCATC TTTTGTCTC
 8221 CAACAGCCAA GTATTTCCA TTAGAGGAGA CTTCCTGTAC ACTTGATGGA TGCTCATTCC
 8281 AAGGTGACTT GGGGCAGTCA GTACAGACTT GGGATGACCT CTGACAGGCT AACCTCTCCC

Figure 23 (continued):

8341 CAACAAGGGC CCTCTATGTT TGCTATGTA TGTAATGTCA GACATTGTCA GGAGTGTCCG
 8401 CAGCACAGCC TGCCCAAGTGT GACGGCTCTC ATAGGTTTCC CACTGTCTTA TCTACACAGG
 8461 GATAACGAGG AGGTAAGCTG CAGTTCCCAG TCTCACTTCA CAGAGGAAGA GATAACCCCA
 8521 TCCCAGGTCA TGAGGCCAGC AGTGGAAAGA ATGAGGATTT GAACTCAGGT CTTCCAAGTC
 8581 CCATTGATAG CATCTCCTCA CAAGTCCCTT GCCACCCCTCA CGATGCCCTTA GACACTTGCC
 8641 TGCCCTTTAT ACTAAGGAGA TGCAAGGTACA AGGGTTTAC CCAATGTAGCA GCTGAGGCAG
 8701 CTGGGGATAG ATACCAGCAG CAGGCCGTAT GTCACCACTC TAACTCCAGC ATCCCCAGTC
 8761 TGTGTTCTG GAGTGTGAAA ATCCCTACTT AACAAAGATTG TGCAACAGTC CTTGGCTCTG
 8821 TGACCCATAG CTGGAAACAG GATTCTCATT GATTGTGGA ACATGGTGGC AGCCAGCCAA
 8881 AAAGAGGGTC TGCATACAGA AGACACGTGT GGCAAGGCCA CAGCAAGACTC TGACTACCTT
 8941 AGCTTACAGA ATTACAAGGT CATAATGTCC TCTGCTTTGG TCACCTCATG TTAAGGACAG
 9001 GCCCCTAATGA AGATGGGGCA GAAGACTGAA GGAATGGCCA ACCAATAACT GGCCCAACTT
 9061 GAGACCCATC CTACAGGCAA GCATCAATTG CTGACACTAC TAATGATACT CTGTTATGCT
 9121 TGCAGACAGA AGCCTAGCAT AACTATCCTC CGAGAGGTCC ACCCAGCAAC TGACTGAAAC
 9181 AGAAAAAAGAT ATCCACAGGC AAACAGTGGA TGGAGGTCAG GGACTATTAT GGGAGAGCTG
 9241 TGGGAAGGAT TAAAAACCCCT GAAGGGGATA GGAACCCCCAC AGGAAGACCA ACAGAGTCAA
 9301 CTAAGAGACC TGTGGGAGCT CTCAGAGACT GAGGCCACAA CCAAAGAGCA TACACAGGCC
 9361 GGTCCGAGGC ACCTGGCAAG TGTGAAGCAG ACATGCGACT CAGTCTCCAT GTAGGGTCTC
 9421 CAATAAGCGG TAGCCTGACT GCAGTATCCA ATCCCCAACAA GGGCTGCATA GTCTGGCTC
 9481 AGTGGGGAG GATGCCCTA ATCCTGCAGA GACTTGTGAG GTGGAGAGCT ATCCAGGGGG
 9541 AACCCACCCCT CTCTGAGAAG GGAATGGGA TGGGGGAGGG ACTCTGTGAA GAGGGGACAA
 9601 GGACAAACAA GAACCTCAAA TAGGTCAAGGC CCTTAAAGGCT TGCTAAGTAG CAGGGCCCA
 9661 GCTCTGCTCT GTTCCCTCAGC CCAAGGTCA GCTCCCAACCT GTTTCTGTGT TTTTCTGGCT
 9721 TTTCATGGGC CTAGGACTTGT GTGACCAGTT CAAACAAATGG GGCCTGTGGA AGACACAATA
 9781 TACAAGACTA GGGACATTCC TGTTCTGCTG ACTATCCATA GCCTGATGTA GGTGGAAGGA
 9841 CCCAATCACT GGATTTCTAC CCTTGCACAA CCTTGACAGC TGAGGGCTC TCAGAAACCT
 9901 ATTTCTCCA CTGAAAAATG AGACTCTCAA ATGAAAGTCG TGACAATCAT CAGGCTTATT
 9961 AAAGAGGTGT ATCTAACCTG AATGGCAAGC AGACAGCAGG CAAATGCTG TATCAACCTC
 10021 TAGGAAGGAC AAAAAGCTGCT CACTGCTGCC CCCCAGGAGG CCATTTGCTG AAACAGCTGC
 10081 TCTCCTGCTG GTGCACAGGC CCTGCCCTCT CATTGCGGCC ACAGCCCCCT CCTGCTGAA
 10141 CCTCCTGTCA GGTCACTGGG AAACAGATCA AGATGGAACA GGACAGCTCC TGATGGTAAA
 10201 TAAAAAACAG TGGTCATGGC TATTGATAGG GGTGTTATGCT TCTTCAGTCC ACACTGTGAA
 10261 GAGCTGTGGG CATGAACACC AGTGTTCGAG GTAGAGTTGG GGTGTTGAAA TTCACAGTGG
 10321 GGTGAGCTCA GTAAATGTGA GCTGGAGGTC ACTCGTGAGA CACACAGTCC TGCTGCTCT
 10381 GTTCCAATA CCTTGAGGAG ACGACACATC TACTTTGTTG AGAGGCCACA GTCTAGTTGA
 10441 CCTGAGAGTT ACCAGTTTCT TATTGTTG TGTTGTTG TGTTGTTG TGTTGTTG
 10501 TGTGTTCTGT GTGTGAGTGC AGGTGCAACAT ATGATAGCGT ACACGTTGAG GTCAAGGAGAT
 10561 AACTATCAGG CGTTGTCCCC TCCTACTTT CCTCGGACTC TGGAGAACAA ACATGGTCC
 10621 TTATTCAGG GGAGCAAGTC GCTGTTGGCT GACACATCTT GCTCACATAC ATTTTACCTA
 10681 GACAATGGAG CCTCCATCAG AGTATTACTT TAGCTCTCA CCGATGGCAA TGCACCCACCT
 10741 CTCTACCCAC ATAGGAGTTG GGTCTCCACA CACCCCCACA CCCCCCTCAC CAAAACGTTT
 10801 TCAGTTACTT TATCTGGTAA AGTTCATCAG AGAATGAAGC CAGTATTAAG AACATGGAAT
 10861 CATTGGGAA CCTGGATCTA GCAATACCCC ACCCTAGATG GAGTTGCTGA GTTTTCACCT
 10921 CAGATTATAA TTCCCCCTCA GCTTCTATGG TTTTATTCTGA AACCAAGGGGA ACTCGATTCC
 10981 TCCCCTTGGG CCACAGACAT CCTGGCTTGT GAATTACAT GTCAATCTACT GCTAATCCAT
 11041 TGGTAGTATG TGGCTCACAG AGACACACTA CAGTCATGGC CAATGTCAAG GTAGGACAGA
 11101 TGTGAATCAT TCCCCCAGTC CTGCTTTT CATGACTAAC CCTCCCTCAGC ACAGTGACCA
 11161 TGAACCTACT TTTCCCCCTCC TTTTATTTTT AGAATTGCTG GAATTTCATA TTTTGAGAAAA
 11221 TAATAGCCTT GGGCAGCATT AAACAAAATC ATCTAGAAAG CTGGTTIAAA ATACAGATGG
 11281 TTGAGTCAGT GAAAGAGTGA GGAATGTCAT TATTGGCCCC TCACAGAGGC TGGCTCACTC
 11341 CAGCAGAGGT GGTGAAAGCT CTTGGACACG GGTCAGGTGC ATAGGAAAGG TNGTCTGGGA
 11401 CACTGAGAAC CACAATTGAA CAAACAGAAC TGTGTTGCTTT TTTTTTTTA AATGAGTTCT
 11461 CAAAAAAATGA CTGGCTAGCT TAGGCAAATA CTTGAGCCA ACCCAACAGA ACATCTTCC
 11521 ATTGATTCTAT TCTGGATCTT CTTTCTAGAC AATACTGAAC TGACCCCTTG TTGGCAGTCT
 11581 CAAGTTTGAC AACATAGGGC TTTGAACTTG GCACAAGGTC CATCACTGTC ACCCAAGCAT
 11641 CCTGGGTGAC CTTTGGGTG GAATATCTTG GCTAACCTTA GATATTCTT TTGGAGTATC
 11701 TTTAGAACAT CCAGGAAATA GGGCTTGATT CTCATCCTGG GACCACAATA TAAGTCACCC
 11761 TAGAACATCCA GGAGATCGTG CAGAGAAACA AGGATCTCTC TCGTGTGCAC CTTCTTCAA

Figure 23 (continued):

11821 AGCAGTGAGT AGTGACTCCA CTAAGACTGAG TTCCCACATCTG AGAGTCCACA GGAGGGCTTIG
 11881 GGGCAAGAAG CAGAGGGAAAG GCACTGTTTG TGTGTTGAAAG GTTTTGACTC TAACAAATTT
 11941 GAAGACATAG ATGACATTGT GTCAGACTAA CAACAACCTA GACTCATGTG GGTTCTGTGTT
 12001 AGGGATCAGA TTTTATTCTAT CAATGACTTG TCTTAGTGTAG TAGAGAAAAGG CTTCTACTG
 12061 GAGTGTAGGC TCAATAATGA CAGAAGAGAT AGCTATTCTC CCTAGGGACT GTGCTGCTCC
 12121 AAGTTGGTG GAGAAAGGCA GTGGGGAAACC TAGATGTGCT CTCTGGGAG GGGGTCTGAA
 12181 GCTGGCTTCAGA TAGAAGGTGT GAAGTTTGCT TGAAACATCT AAACAGAATT ATAGCTTAGG
 12241 AAAGTGAGCA GGCAGGGAG GGAATGTGTT GCATATGTAT ATGTACATGA ATATATTATG
 12301 TTATAGATAC ACACACATTG GAACCTCATT TGCAGATGAC AGAAAATAGG TTATTTGCC
 12361 TCTCTTAACG GCTAAGCACA ATGACTTCCA GTTCCATCCA TTTCCTGAAA TGCCACAATT
 12421 TCATTTTCA TTGTTGGCTGA ATAAAAATTCC ATTGCAACT GGGCCCTACT TCATCCACTC
 12481 CTGAGGGCAG GCATATCCCC TGGCTCCATT TCTTACCTAT TGTGAAGAGA AGTGCACATG
 12541 TCTTGTGAA AGGCAAGCGT GAGAGAGGCA GGCACATAATT GTGGGTTTT GTTTCTCTT
 12601 CCTGCTATGA CTCTCCATTG TGCAGAAACCA AAGATCGATA AAAGCCGCCA CCATGAAAGC
 12661 CATCTTAATC CCATTTTAT CTCTTCTGAT TCCGTTAACCG CCGCAATCTG CATTGCTCA
 12721 GAGTGAAGCCG GAGCTGAAGC TGGAAAGTGT GGTGATTGTC AGTCGTATG GTGTCGCTGC
 12781 TCCAACCAAG GCCACGCAAC TGATGCAGGA TGTCAACCCCA GACGCATGGC CAACCTGGCC
 12841 GGTAAAACGT GGTGGCTGA CACCGCCGG TGTGAGCTA ATCCGCTATC TCGGACATTA
 12901 CCAACGCCAG CGTCTGGTAG CCGACGGATT GCTGGCGAA AAGGGCTGCC CGCAGTCTGG
 12961 TCAGGTCGCG ATTATTGCTG ATGTCGACGA GCGTACCCGT AAAACAGGGG AAGCCTTCGC
 13021 CGCCGGCTG GCACCTGACT GTGCAATAAC CGTACATACC CAGGCAGATA CGTCCAGTCC
 13081 CGATCCGTTA TTAAATCCTC TAAAAACTGG CGTTTGCCTA CTGGATAACG CGAACGTGAC
 13141 TGACCGGATC CTCAGCAGGG CAGGAGGGTC AATTGCTGAC TTACCGGGC ATCGCAAAC
 13201 GGCCTTTCGC GAACCTGGAAAC GGGTGCTTAA TTTCCTGCAA TCAAACCTGT GCCTAAACG
 13261 TGAGAAACAG GACGAAAGCT GTTCATTAAC GCAGGCAATT CCATCGGAAAC TCAAGGTGAG
 13321 CGCCGACAAT GTCTCATTAA CGGGTGGCGT AAGCCTCGCA TCAATGCTGA CGGAGATAATT
 13381 TCTCCCTGCAA CAACGACAGG GAATGCCGG ACGGGGGTGG GGAAGGATCA CCGAATCACA
 13441 CCAGTGAAC ACCTTGCTAA GTTTCGATCAA CGCCCAATT TATTGCTAC AACGCAAGCC
 13501 AGAGGTTGCC CGCAGCCGGC CCACCCCGTT ATTAGATTG ATCAAGACAG CGTTGACGCC
 13561 CCATCCACCG CAAAAACAGG CGTATGGTGT GACATTACCC ACTTCAGTGC TGTTTATCGC
 13621 CGGACACGAT ACTAATCTGG CAAATCTCGG CGGCCACTG GAGCTCAACT GGACGCTTCC
 13681 CGGTCAGCCG GATAACACGC CGCCAGGTGG TGAACCTGGTG TTGAAACGCT GGCGTCGGCT
 13741 AAGCGATAAC AGCCAGTGGG TTCAGGTTTC GCTGGCTTC CAGACTTAC AGCAGATGCC
 13801 TGATAAAAACG CGGCTGTCTAT TAAATACGCC GCCCCGGAGAG GTGAAACTGA CCCTGGCAGG
 13861 ATGTGAAGAG CGAAATGCGC AGGGCATGTG TTGGTGGCA GTGTTTACGC AAATCGTGAA
 13921 TGAAGCACGC ATACCCGCTT GCAGTTGTAG AGGTACCCGG GGATCACAAAC TTGCCCCCTG
 13981 AAGAGGAAGA ACAGAAGGAT GCCACAACCTC TCCTGCTGGC TACTCTCCAG TGGTTTCATC
 14041 TTACTCTGA TGGCATTTCC CTCTAGAAAG TGCTACTATC ATCCACACAT TTCTACCTGA
 14101 GACCACCAA AGGACCCCTCC CAAATTCTCT TCCCTCTGTGA GTAGTCTCCA CACCTGTTAC
 14161 CACCATCCA GAATTAAAAT CCTAATGCA CTCTGGCTG TGACTTGCT CAGTCCTTGC
 14221 AATAAGAGTT GTTGGCAGTG CCAGGCGTGG TGGCGCACGC CTTCATTCAGC AGCAGTTGGG
 14281 AGGCAGAGGC AGGCGGATTCTGAGTGGCA GGCAGGCTG GTCTACAGAG TGAGTTCCAG
 14341 GACAGCCAGG GCTATACAGA GAAACCTGT GTCGAAAAAC CAAAAAAA AAAAAAAGTT
 14401 GTTGGCAGAG GTTGGTTTAT ATACCAGGTG GAGATTTCAA ATGAGTGGCT GAAGCTGTAG
 14461 CCAGAAGGAA CTTAGAGGAT AGCTCATAAC TAAAAAGAA ATGTAGAGAG TAGCAGAAAC
 14521 ATTGAGAGAG TGGGCACACA GCCACTGTGT GAATGTGGCA GAACACAATC CAGCAGCTA
 14581 TACATGCATA AGTGTATAATT GGCGCCATCC TGACTGATGA GACACAGGAA AACAGATAGA
 14641 CGGGGTTAGG TGGCCATGGC CTTTCTGCT TGCCCTCTCC TAAGGGTCAT CTCAAGACCT
 14701 TATGCTCTCT TAATCTCTCC ATTGCTACTT AGCTCTAGA TATCACCTCC AGATTAGTCT
 14761 CCTGGGTAC ATCAGTGATC CTGGTGTAT CCAGGGCTTC CTGATTCCAT CTTTGTCTATA
 14821 GAGGCTGCAA CTAAGAGGAT CTTCTTAACTA CTTCACACCC TGATGCCAAA AGGAAGACAC
 14881 AGAAGTTCAC AGAGGTGAAG TGATTCTGT AGGACATACA GTGAGCAAGC ATCAGGGTCC
 14941 GGATTATCTG ACTCTACTCT AACTTTATG TAAATGTGCT TTATGCCATT AACACTGTCA
 15001 TTCCCTGTGCT TCAGCTCTGG GAGACTCCCA AGCCTCTTA GGCACAAGCC ACAATTAAGG
 15061 GACTCTGACA CTCCTGCTTG ATTAATTAGC ATGGTGGCT CTATGTTCC AGATTCTATG
 15121 TTGTTTCACT TTCCATATAG GCTATGAAGG GTGTGAGGAA ATTTTTGGG GACAGAATTG
 15181 GAGGCAATCC ACCTCTCTCA GGAAGCCTCT ATCTGGAAAA GCTTACAAC CAGGGACAGT
 15241 AACTGTAGGC CCAGCTCTGG GTGTCCAAAA TGGTTTTAT GGTTGAAATC TGCAAAGCCT

Figure 23 (continued):

15301 TCCATGTGCT CAAAGGTTG AACATGGAGC CTCCTCCTGG TAACACTGTA TTGGAGGGCTT
 15361 TTGAGACTGG ATGCTCTTGTG GTCCCATGTT TTGCTACATC ATCTGTCAAG ATATGACCCA
 15421 GGCATGCTAC CAGCTACCAC AGACTATGCC TCTCCAGCTT TCATGTTCTC CCCACCATGA
 15481 TAGACTTGTA TCTCTTAAAA ATGGAATCAA AGCAAACCTT TCCTGCATTA AGTTTTTTT
 15541 TTTCTGTTAA GTGTTGGTC ACAGGGACAA GAAAACACTC AATACAGATA ATTAGTACCA
 15601 GAGTTGAGGT TCATTGCTCT AGCAAGTTGG ATCAAATTTT TAGGGCTTGTG GAAGTGATTT
 15661 ATAAGAGACA TGTAGAAGAG TCTGAAGCTG TGGGCTACAG AAGTGTCAAC AGTTTTAAG
 15721 AATAGTTAA TACACCATGG GAATTGTGAA AATCAGAATG CTCACACAAA GGCAGACAGG
 15781 AAAACGTGAG CATGTGGCGT GTGAGAGGGC ATAAGAAGGA ACCTAGGGGG AAATGAGCTA
 15841 GAAGCCATTC GGCTACGTTA GGGAACGTGT GTGGCTGTGC TTGGCCCATG CCCTGCAAT
 15901 CTGAATGAGG CCAAATTTTA AAGGAGTGGG CTAACTCGAT TGTCAGAGAA AATATCAAGA
 15961 CAGACCACCA CTCAGGCTAT GCCGTGTTTG TGACCGACCA GCTACTCTTA GCCAGCTCTA
 16021 TTGTGAAATT CCAGAGCAAT TATCAGAGCA TGAAGATACA TACAGTTTAG TGAAGTAAGG
 16081 GGTGTGGTC CCTAAGTGGG TGGTGCATAA ATCTATGTAG GTGATGCCATA AGTGACACTT
 16141 GATAATCCAA AATATCAGCA ATGTGGAATG TCTTCCAAGG AGACCTGTAG ACACACATT
 16201 TAGAACTTTG CTCATGGCTG TAATAAAATAG CTAGCTAGAA ATCATTCCCT GAAGAGGTTA
 16261 GTCTGAGTTA CGGTCCAGG GCAAACATTC AGTGATGGCA AGGAAGGCAT TGCAGTCAGG
 16321 AGCCAAAGGT CAGCTGGTCA CATTGCATCA AGAGTAGAGA GTCAAGAGTGT GAGTAGAAAG
 16381 AGGATACAGG TTATAAAACC TCACTGTCCA CTCTCAGCAA TCCATTTCCT CCTAAAAGGC
 16441 TTTACCTTCT AAAGATTTTA GTCTTCAAA CCAGTACCCAG TAGCCTGGGA ACAAAGTTG
 16501 AAACAAATGA GCCTTTGTGG GGCATTTCAC ACTTAAAACA GGGCATCACC TAGGAGGAGC
 16561 CCTGTGTGCA GTAGGAAGTG TGGCCTCTGT GTCAGGAATG CTCAGGCTAA TAAGGGGTCC
 16621 TCTATCTGAG GGACCCATATG AAGATTCAAC AAGTAGTTGT GAGAATTCCC TGTAATGGG
 16681 TGCTACCAAT TTGACATTG TAGACCTGCT ATTGTGTGCT TCTTATTGG GCTCTCCAT
 16741 CTCCCAACTT TCCAACCCAT ATTCCACATT AATCCCTTCC ACCACCATGC AACACTAGGT
 16801 AGGAGAGAAG GAAGGTAGA AGAGAAAGTG GTTATAGATC ATTATAGACT ACTTCTGCT
 16861 GATTAGGGC AAGTCCAATC GTCATTGTC GGATACCTCC ACCAGCAAC CAGCAAACCA
 16921 GCAAATCAGA AACAGAAAAA GCAGCCAACA AGGCAGCACT ACCAGCAGG ATTGGGGTCG
 16981 GTAGCGTGGG AGCAGTCACT ACTGGCTTC TCATGGCTTT GGCATTAAATA CTCTCTCAAG
 17041 AAATTCCGTA ATTTCCTCCC CACCACTGA AATTCCTGAA TTTTAAATGC AACTATCTA
 17101 CAGCTGGCAA AAATCACATC TCTCCTAGAG CACAAGACAA ATCATAGTTA CTGGCTATT
 17161 GCAATCTGAA GCATCTCAAT ATCCCACACC TGGGATTAAA ACAAACAT ATTCAACATCA
 17221 CATAACTGTT TTTTTTCC AATTTTTTAT TAGGTATTTT CTTTATTTAC ATTTCAAATG
 17281 CTATCCCCAA AGTCCCCAT ACCCTCCAC CTCCCTGCTC CCCTACACAC CCACTCCAC
 17341 TTTTGACCC TGGAGTTCCC CGGTACTGGG GCATATAAG TTTGCAAGAC CAAGGGGCCT
 17401 CTCTTCCCAG TGATGGCCGA CTAAGCCATC TTCTGCTACA TATGCAAGATA GAGACACGAG
 17461 CTCTGGGGGT ACTAGTTAGT TCATATTGTT GTTCCACCTA TAGGGTCCGA GACCCCTTCA
 17521 GCTCTTGGG TACTTTGTCT AGCTCCCTCA CTGGGGGCTC TGTGTTTTAT CTAATAGATG
 17581 ACTGTGAGCA TCCACTTCTG TATTTGACAG GCACTGGCCT AGCGTCACAT GAGCCAGCTA
 17641 TATCAGGGTC CTTTCAGCAA AACCTTGCTG GCATGTGCAA TAGTGTCTGC GTTGTGGGT
 17701 TGATTATGGG ATGGATCCAC TAGTTCTAGA GC

//

SEQUENCE LISTING

<110> University of Guelph
Forsberg, Cecil W.
Golovan, Sergeui
Phillips, John P.

<120> Transgenic Animals Expressing Salivary Proteins

<130> U Guelph - 03

<140>
<141>

<150> 60/130,508

<151> 1999-04-23

<160> 7

<170> PatentIn Ver. 2.1

<210> 1
<211> 20623
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Lama2/APPA
plasmid

<400> 1

tcgagagtat ctttgtcagc tggcctcca acaaagggtt actgttgcac 60
atctaaacta attaataat ccctcaccgg caaatcttc agtcaactaag ttagcacat 120
tggtaacaa gttctccaaa ggagagatac agatgagtgc gtatagggtg gacctggctg 180
ctgaggagac acctgcattt gactaagaag agccacgggtt ttagttaat ggtgtggagt 240
agggtgggtc tggggacag tagaaaatcg agaggcatgt gccgttttagt gaactgtatgg 300
aagctacccc aaacgacaga gattgtcagt caggccaatc cggttcgagt ttgatggca 360
gcccggacagt gagacagaca cacctactca gttggaggaa ggatgagaac aatggccagc 420
agggatttag agaccctgac aggcgcacagg ccctaaacaca cacacctacc acctcaactt 480
acaaagctgc caaagaccaa agacttgc tccattagaa atgacagctg gcttgacccg 540
acagcataat aagcagagtg tactctgatt ggagaacttt aatgtgttcc attcagtatt 600
ataaaaaggac agtattacag attttgttgc acactgctgt tacatgtggg gcagtgtgtc 660
tttaagttagg gtaaaagtact ctttaaaaat gggccctaga tatttttcc tttaactcaa 720
gtctcttact gtttaatga tttttatgtt gtttaatatg gggaaaaag aagcgtaaat 780
ggacaatata tatttagaga aagatggta gctgtcagaa aaatatgcaa atcaaaatca 840
caccaagact gcagcacacc cctgtcagat ggctgtgtc aagaaaataa atgacaatga 900
gtgggtgtga agatgtacta aaggaaaca cacacacaca cacacacaca 960
cacactggag caaccactgt ggaaatcagt atgaatggtc ctcaaaaacc tgaagataga 1020

gcggggcgtg gtggcataca cttttattcc cagcactggg gaggcagagg caggtggatc 1080
tctgagttcc aggccagcct ggtctatagc acaggttcta ggacagccag ggctacacag 1140
aaaaaccctg ccttattaa accaaaccaa accaaaccaa accaaaccaa accaaaccaa 1200
accaaaccaa accaaaccag accaaaccaa aacactgaag atagaacttc agtattccat 1260
tccttagat ataccaatg gagactaagt cagcaagaca cctgcacagc catgttca 1320
actacactgt tcaccacagc caggctgtgg aaccagcctg agtgcctatg ataaatgaat 1380
ggataggtaa ctttcaaggt aaatggactc tgctgtgtac atgcctcaca ttctgtttat 1440
tcattttctt ttaggatgtc tccattcagg agtcacatgg tagttctatt ttca 1500
tgaagatact acactggtcc ccacagtttta cactttatc agcagtgaat aagggttcct 1560
ctatcattac catcattgt tgtaattttt ctgtatgacc ctcttctga cagggataagg 1620
atgtatatac agtgcgttgg agtacaactt gtttctaag tattttatgg ccccttgcatt 1680
ttcttctttt gaaaactgtc ggttcctgac atctgctcag gtattcattt gatgttgc 1740
cttgggttt tgagttctta tgaattcttag atgttaatc cctgcctgtg gttctctccc 1800
attctgttagg ctgcctcctc accctggcaa ttgttgcct tggttgcag aaactttga 1860
cttcattgaa tctcatttgc cagtttccc tccctgtca tagcctgagc taatgcactg 1920
gttttacag agccctggtc tatgccttta tccctctctg gcagcttcgg agtttcat 1980
cttacattta gatctttgat ccactttgaa caagtttgg agcagggtga gagatacga 2040
tctagttcca ttcttcata tgcgtatcata gtttacatag catcgttgc tgaagagg 2100
ttatatttattt tttaaataat gtgtcataaaa aaacgagggtg gtttagcag tgcgttgc 2160
tttcttgc tttgtatcata caggtcttgc tttgtgtcag tctcatgatg ttttattgtc 2220
atggctctgt catacagtct gaggtcaggat attgtgatatac accttcgatg ttgccttc 2280
agactcaggat ttgccttggc caggagtcat cttactcagt gctcttagag ctcccccagc 2340
atgttagctgc tactattctt agttgataaa tcaggaaact gggctcaga gagattaact 2400
gtcttgcact acttctgggg aggtgaaacg tggagacact aaactgtt gatgttgc 2460
tgctccagta gctgtcgggt gctggctac agcaaaagcac ctatactata tattactcag 2520
gaggtggaaa aactcagect cccttgggt tcccaagctc ccaggtgtcc agtcaactg 2580
ggaaacctca tggagtcata aaggaagggt tgaggatcata tggggcagcg atgaggagcc 2640
tggggctggg atctccaaa cacctggata tccagatgcc acttgggtcag ggggagttgg 2700
gaacagagtt gggatgtcca tggacctgtg acaaggccag ggccagggggg aggataactc 2760
tggcttact aatttgcgaa agtccttago ttagcagcag ttgttgcggc gcacagaggg 2820
gccttctgtc agaggctcag gcagtgcgc tctgttagcag aaggctttctt ccatgttccc 2880
catgggtttt cttgtatgaaa gagacagtcc ttggctccaa actgggtttat tgattgttca 2940
ttgtggaaaaa tgggtgcaca ccaccttctc aggggtggacc agagatcaaa tacctttgc 3000
agggaggaat atctggaaag ggacgcttac tggctaaacc ctcaggccct ctagatacat 3060
cattagcatg gagaactctg ttctggctca catgaccaca ggccacat 3120
catgtggaaa gtgtggcaca ttgttcttagc caggaatctg ttagggagcg tggagccacc 3180
taccatccca ggtgggtgcc tgggtgccag ggaccctgaa cccgctcaac ttaccaact 3240
ttcctggcag ggtccactgt cctacacaga agctggagga ggtgtgaggg ttgttgc 3300
gtggaaatgtc ccatgtctgc tggggctcag tttctccacc tgcgtatcata tgggttgc 3360
ataaaaaatgt gggataactttt attattctct gactcgatc tggggggggc gatcggtgc 3420
agtccaggaa ccacaccctg aggttctgc actgaaggga ctccctaaatg ctctggagtc 3480
tctcccttc acagagctgc caaagtcttag gttctttga ggataacaga gccatgttc 3540
gtaaggcagac aacagcattt gtttactcaa ctttcttttgc tca 3600
aagttgagac accatgctgg cttgaggaag acttctaaag ccagacaaact gtgcaggaa 3660
gaagaagaag gggcaagtgg agttagcctg gatgtagccc tcaaaatgtc cagagaccag 3720
ccatgaaggc tcaagtggag ggcaagaccc gcagcagccaa agcatctggc aggagaggat 3780
cctggaaacc ccttaccat gacacacattt cttctgtcag gtcacactta ataggccatt 3840
tcttatttttgg atctatcatg gtgttctgtg cgagatataat gaggtgtttat gtcgcgaaca 3900

aaaatgttata taaaacaag tccccccccc ttgtcactgc tgctaagaat gtagcagaaa 3960
ttgtctcaag tgtctctcta atcagaaaca ataaaggctc cttggattc aagccctcca 4020
gttccctct tccttgcga gccttggaca cccatacaaa ctcctggat gctacagtc 4080
tggcagaga ctccaagggtc gggagagact gatggtacaa aagcaaaata cttgttggg 4140
ggtacaccca ctccctgcct tgcgtggcct ctgcagtcag tcctgcagac aggcctcag 4200
tgggtcttcc atggcaaca cgcagggga ggcaatggat gggataacc acaccctgg 4260
tagtttaccc cggccatgct ctctgcttt catccctctt ctgcctctg ccacggctt 4320
ctctgcagga atcatatctt catattggcc cacaggtgtt ctcctcaccc tagctatgt 4380
gtttacttta gagtgaccc agcagggtc gtggaatga gttctagaag gtcacggag 4440
atgctaggta agaaacgtct tctaactact gaggtacta agttcctgtt gttgtctct 4500
gcctttccct tggtaaagtc accttgaagt tagtgcagaa gaaatcagag cccagtcaca 4560
gagtaaatat ggtcctgaag atttccttg agtgcggcaga atccatgaca tttcaagagc 4620
cctcttgc tccatagtca ttgggggtt tatctctgc ttgatgtatg tttgtgtt 4680
tatcaaagag ttagatgggtt acataagagg tgctctaaag gacagagagg atttgcaatt 4740
gtggcatgtg acatcctcag gccttgcct ggtgccagga ggaactgtatg cagaaaagag 4800
taagaggc tttctggag gtcgtacta tagaggagat cttacagtgc attccctcct 4860
ccaggccctg cctgaggata gacatgtct gactgcaact gaaacagagg ctgggatgg 4920
agagtttagt tcacagaagg gagggtggg gatggatgtc tgctgggttc tgggtctcat 4980
caccagctcc tgaccacccg gtcagcccat gtgcatttcc catagcttcc ttttgcatt 5040
tttactcagt gtgggttttgc ttgggaccca gcagaagcca gtcccaggct gacagctgt 5100
gatacacagg gcacgtatg ggtcctcagc ctgaagcagt caggctggca gaagagaaaag 5160
accagcacac attccttcaa ccaactatgt cttgaaaaac aaacatattat tatcacat 5220
attgcatttgc ttagacagct aaaatgtact cgggttagcat gactccagggt ggggatatct 5280
gcaagtgc ttagtggcag agggacagcc aatgtgaggc aagaaggaat tctggctcaa 5340
cacagcttag ctccctgggtt ttggttccaa ctttgagat ttgaccacaa gcaactttatt 5400
tttgacat taaaacagag cacaactttg gggaaaagtt ttcttatgaa aattatcaca 5460
ataaaagcttgc tggatgtact acattaaaat gccttgcac agtataatgtc ccctcttcc 5520
caagaatggc tctattgtact gagaataat gttcaggata aagatccagg aagaaaagat 5580
cagggataag taaaatacta aactcttttgc caaagtacat agacccttcc tcataacaat 5640
gggttctatt gactgacaag cactgctcag gagttgggaa agagtcttagc ataagcacga 5700
tagcctggag actcttagtgc ggtcttagtct tacagacagc aaaaatcacc aggttacaaa 5760
ctacattcat ttccagttt ctgatcaggc acaggtatga atcccttgc ttgaagagaa 5820
aagtccatgt gttaaaata tctggtttctt ccagtctat tagcgagaag acttgagccc 5880
tatacaactc ccacccctggag tgacatcctg tcttcattgtt atattacata cctagacacg 5940
ctcatctcac agacttagga ctttgcatttgc tgatctccat ttctgtatccc acttccaccc 6000
ttgccttgat agtgcattt tcttcactgc cttgggtgaca accatgttat cctctgtgt 6060
tttgagtttgc accatccat gatccatgtt gatgcacca gtcacacagtc tttgtcttcc 6120
tgtctggatg catgtcaatc tctacacaac aacccttccc cgtcactcag atcttcctcc 6180
attaacacat acatgggtc gaaaggacta gggagcttcc cttcagttggg gagctagctg 6240
gctattggc cttttgcact gtcaggaaag gcccccaatt gctgagacaa gaacttagat 6300
tcttcattat tgactctaac tcatgtatca agcagaagct aatgaatagt tatcaacagg 6360
atcagagggtt ccagtgtaaag acactttgac atgaaaagaac ggaggaagga cagatggatg 6420
cataaaagca ggaccactgc cccaggaagg tcctggaaac tgatgcaggc caaaggacag 6480
gttataaaacc aaatcttagg gagtcaggaa gagcacagag gagctcaacc aactgaccac 6540
tgcttaggg ctaccaaccc aatccctccct gtggaaacag ctaagctatc agccaaagggt 6600
aataaaacagg caggacactgt ggatgacatg gagagcatag ggaccctggg tccagcctt 6660
agcaccctgca ctctcaggat actccacccat tgcgtttag agagccttagg gataactgggt 6720
ccagccttgc gtagccatcact tctcagggtt cccatcaact gttgtttggaa gaggcttaggc 6780

accctgggtc cagccttcag tacctgcgtc ctcaggacac cccaccattg tctcttgccc 6840
cgtctttctt tcctcttctt cccttcatt gtctttctc tgtttcttc ttgactctcc 6900
tttccccctca caccctcaact ctatgttctcc cttccctctt ctgcattcacc ctattcttc 6960
tgtggccctt ccactttctt ttatctctca tgcttcttc ctccctcaaa tacttgtcac 7020
ccactatact tcaggggcca gctcttagtga caaagctgtt aatagcaaga ctctcagatc 7080
tccaacggct cagaggagcc agacccacca agaactctct ccaggtccaa tttcaggttc 7140
cttcgaaaggc tttcagcaaa tgctcaggga acatgccact aacaagaaga tgcaaattcc 7200
agttgagagt gggaaaggcc ctgcgttagg tcccatcttc caggccaagg tcagagggc 7260
tctgtgtaat ccggattgac agggctcaga acaatgttt gttttaagg ttatatttatt 7320
ttaggtgtta gtgtcttgc ttgcattgacc ttatgtgcat catgtgtgtg caggttcctg 7380
atgacagtag aggagggctt tgaatccctg gggataggaa gttacaggaa attataagct 7440
gctttgtggg tcttctagct ttcccaacag aagtgaatgc tcttcaccac tgagccatct 7500
ctcttagggccc aagagacatt gctttatgga tataattgtg tgtgtgtgtc aacattgagg 7560
aaaggaaat aaaaaaaaaa cttcagccgc taagggtgtc cagtttcaact aattgtact 7620
tttagttgtg ataaaatggc aggtgcttca acatttatata atacaaaaac ttccctgtc 7680
gtgggtcaac tgtgagaact gggtaagtg ggtgaggctt ctttttctgt ctctgtctct 7740
gtctctctcc ttccatttctt tcttaaagga aataaacatt gcagctgggt tatagtctcat 7800
caatatggaa gttacagaag taaaaaaagg cattgcctt gttgggtggg ttaccagctg 7860
atttttgggt gtctgtcaag gaggtctggg gactggctgc tctgtctctg tctgtatgag 7920
tgagggaaagt ctggggagca gattccctaa ctttcagcct ggctgggtc ctgagtgaac 7980
ccagcctctc tggccttagt agcttttcc aaacaggaat ctgagtgggt acagggaaaca 8040
agtaccagcc cattgtttaa gtcccgagggt tagtgaggggc aggaagctgc catagctggg 8100
attagtagtt gtattggatg taggaagtcc tattctggg cagctaattcc ttaatgttcc 8160
actggagatt ttcaatgaga aatttataccc acggccata tggcccccattt cttttgtctc 8220
caacagccaa gtatttcca ttagaggaga ttccctgtac acttgatgga tgctcattcc 8280
aaggtgactt gggcagtc gtacagactt gggatgacct ctgacagcct aacctctccc 8340
caacaagggc cctctatgtt tgctatgtaa tgtaatgtca gacattgtca ggagtgtccg 8400
cagcacagcc tggccctgtt gagggtctc ataggttcc cactgtctta tctacacagg 8460
gataacgagg aggtaaagctg cagttccctag tctcacttca cagaggaaga gataacccca 8520
tcccagggtca tggccctgtt agtggaaaga atgaggattt gaaactcagggt ctccaagtc 8580
ccattgtatag catctcttc caagtccctt gccaccctca cgatgcctta gacacttgcc 8640
tgccctttat actaaggaga tgcaggtaa aggggttac ccatgttagca gctgaggcag 8700
ctggggatag ataccagcg caggcctgtat gtcaccactc taactccaggc atccccagtc 8760
tgtgttcctt gagtgtgaaa atccctactt aacaagattt tgcaacagtc cttggctctg 8820
tgaccctatag ctggaaacag gattcttattt gattttgttca acatgggtggc agccagccaa 8880
aaagagggtc tgcatacaga agacacgtgt ggcaaggcca cagcagactc tgactacatt 8940
agcttacaga attacaaggt cataatgtcc tctgttttttgc tcaacttcatg ttaaggacag 9000
gccctaatga agatggggca gaagactgaa ggaatggca accaataact ggcccaactt 9060
gagacccatc ctacaggcaaa gcatcaattt ctgacactac taatgtatct ctgttatgt 9120
tgcagacaga agccttagcat aactatcctc cgagagggtcc acccagcaac tgactgaac 9180
agaaaaagat atccacagggc aaacagtggc tggaggtcag ggactattat gggagagctg 9240
tggaaaggat taaaaaccctt gaaggggata ggaaccccaac aggaagacca acagagtcaa 9300
ctaagagacc tggggagact ctcagagact gagccaccaaa ccaaagagca tacacaggcc 9360
ggtcccgaggc acctggcact tggaggtcag acatgcagct cagtctccat ttaggtctctc 9420
caataagccgg tagcctgact gcaatgttca atccccaaaca ggctgcata gtctggcctc 9480
agtggggag gatccccta atccctgcaga gacttgatga gtggagagct atccagggggg 9540
aaccacccctt ctctgagaag ggaatggggta tggggaggg actctgtgaa gaggggacaa 9600
ggacaaacaa gaaaccttcaaa taggtcaggc cctaaaggct tgctaaaggtag cagtggccca 9660

gctctgtcct gttcctcagc ccaaggctca gctcccaccc gtttctgtgt ttttctggct 9720
tttcatgggc ctaggacttg gtgaccagtt caaacaatgg ggcctgtgga agacacaata 9780
tacaagacta gggacattcc ttttctgtcg actatccata gcctgtatgt ggtggaaagga 9840
cccaatcaact ggatttctac cttgcaccaa cttgacagc tgagggcctc tcagaaacct 9900
atttcttcca ctgaaaaatg agactctcaa atgaacgtcg tgacaatcat caggcttatt 9960
aaagaggtgt atctaaccctg aatggcaagc agacagcagg caaatgtctg tatcaaccc 10020
taggaaggac aagaactgct cactgctgcc ccccaggagg ccatttgcgt aaacagctgc 10080
tctcctgtcg gtgcacaggc cttgccttct cattgcagcc acagcccctt cctgtctgaa 10140
cctcctgtca ggtcactggg aaacagatca agatggaaaca ggacagctcc tgatggtaaa 10200
taaaaaacag tggcatggc tattcatagg gttttagtct ttttcagtcc acactgtgaa 10260
gagctgtggg catgaaccac agtgttcgag gtagagttgg gtttctgaaa ttcacagtgg 10320
ggtgagctca gtaaatgtga gctggaggc actcgtgaga cacacagtcc tgctgttct 10380
gttcccaata tcctgaggag acgacacatc tactttgttc agaggccaca gtctagttga 10440
cctgagagtt accagtttct tatttgcgt tttgtgtgt tttgtgtgt tttgtgtgt 10500
tgggttcgt gtgtgagtgc aggtgcacat atgatagcgt acacgttgag gtcaagaggat 10560
aactatcagg cttttttttt ctttactttt ctttggactc tggagaacaa acatgggtcc 10620
ttattccagg ggagcaagtc gctgttggtc gacacatctt gtcacatata atttaccta 10680
gacaatggag cttccatcg agtattactt tagtcctca ccgtatggcaa tgcaccaccc 10740
ctctaccac ataggagttt ggttccaca caccggcaca cccccccttccac caaaacgttt 10800
tcagttactt tattttttt tattttttt tattttttt tattttttt tattttttt tattttttt 10860
catttggaa ctttggatcta gcaatcccccc acccttagatg gatgttgcgt gtttccaccc 10920
cagattataa ttcccccccta gtttctatgg tttttttt tttttttt tttttttt tttttttt 10980
tcccttttgggaa ccacagacat ctttttttgggaa ccacagacat ctttttttgggaa ccacagacat 11040
tggtagtatg tggctcacag agacacacta ctttttttgggaa ccacagacat ctttttttgggaa ccacagacat 11100
tgtgaatcat ttttttttgggaa ccacagacat ctttttttgggaa ccacagacat ctttttttgggaa ccacagacat 11160
tgaacccactt ttttttttgggaa ccacagacat ctttttttgggaa ccacagacat ctttttttgggaa ccacagacat 11220
taatagccctt gggcagcatt aaacaaaatc atctagaaag ctgggtttaaa atacagatgg 11280
ttttagtcat gaaagagtgtt gaaatgtcat ttttttttgggaa ccacagacat ctttttttgggaa ccacagacat 11340
cagcagaggt ggttgaagct ctttttttgggaa ccacagacat ctttttttgggaa ccacagacat ctttttttgggaa ccacagacat 11400
caactgagaac cacaattgaa caaaacagaac ttttttttgggaa ccacagacat ctttttttgggaa ccacagacat 11460
caaaaaatgtt gggcagcatt aaacaaaatc atctagaaag ctgggtttaaa atacagatgg 11520
atttgcatttcat ttttttttgggaa ccacagacat ctttttttgggaa ccacagacat ctttttttgggaa ccacagacat 11580
caagtttgcatttcat ttttttttgggaa ccacagacat ctttttttgggaa ccacagacat ctttttttgggaa ccacagacat 11640
cctgggttgcatttcat ttttttttgggaa ccacagacat ctttttttgggaa ccacagacat ctttttttgggaa ccacagacat 11700
ttttagaataatgtt gggcagcatt aaacaaaatc atctagaaag ctgggtttaaa atacagatgg 11760
tagaaatccca gggcagcatt aaacaaaatc atctagaaag ctgggtttaaa atacagatgg 11820
agcagtgttgcatttcat ttttttttgggaa ccacagacat ctttttttgggaa ccacagacat ctttttttgggaa ccacagacat 11880
gggcagaag ctttttttgggaa ccacagacat ctttttttgggaa ccacagacat ctttttttgggaa ccacagacat 11940
gaagacatag atgacattgtt gtttgcatttcat ttttttttgggaa ccacagacat ctttttttgggaa ccacagacat 12000
agggttgcatttcat ttttttttgggaa ccacagacat ctttttttgggaa ccacagacat ctttttttgggaa ccacagacat 12060
gagtgttaggc tcaataatgtt gggcagcatt aaacaaaatc atctagaaag ctgggtttaaa atacagatgg 12120
aagtttgggtt gggcagcatt aaacaaaatc atctagaaag ctgggtttaaa atacagatgg 12180
gctgggttgcatttcat ttttttttgggaa ccacagacat ctttttttgggaa ccacagacat ctttttttgggaa ccacagacat 12240
aaagtgttgcatttcat ttttttttgggaa ccacagacat ctttttttgggaa ccacagacat ctttttttgggaa ccacagacat 12300
ttatagataac acacacattt gggcagcatt aaacaaaatc atctagaaag ctgggtttaaa atacagatgg 12360
tctcttaactt gggcagcatt aaacaaaatc atctagaaag ctgggtttaaa atacagatgg 12420
tcatttttca ttttttttgggaa ccacagacat ctttttttgggaa ccacagacat ctttttttgggaa ccacagacat 12480
ctggggcag gggcagcatt aaacaaaatc atctagaaag ctgggtttaaa atacagatgg 12540

tcttgttcaa aggcaagcg tggaggc ggcactaatt gtgggtttt gtttcttctt 12600
cctgctatga ctctccattt gtcagaacca aagatcgata aaagccgcca ccatgaaagc 12660
catcttaatc ccattttat ctcttctgtat tccgttaacc cgcataatctg cattcgctca 12720
gagtgagccg gagctgaagc tggaaagtgt ggtgattgtc agtgcgtatg gtgtgcgtgc 12780
tccaaccaag gccacgcaac tgcgtcagga tgcacccca gacgcattggc caacctggcc 12840
ggtaaaactg ggtggctga caccgcgcgg tggtagctatc atcgccatc tcggacattt 12900
ccaacgcccag cgtctggtag cgcacggatt gctggcggaa aagggtgcgc cgacgtctgg 12960
tcaggtcgctg attattgctg atgtcgacga gctgatccgtt aaaaacaggcg aagccttcgc 13020
cgccggctg gcacactgact gtgcataaac cgtacataacc caggcagata cgtccagtc 13080
cgatccgtta tttaatcctc taaaaactgg cggttgcacca ctggataacg cgaacgtgac 13140
tgacgcgtatc ctcagcaggc caggagggtc aattgtgtac tttaccggc atcggcaaac 13200
ggcgttcgc gaactggAAC ggggtgtttaa tttccgcaaa tcaaaacttgt gccttaaacg 13260
tgagaaacag gacgaaagct gttcattaaac gcaggcatta ccatcgaaac tcaaggtgag 13320
cgccgacaat gtctcattaa cccgtgcggt aagcctcgca tcaatgtgtc cggagatatt 13380
tctcctgcaaa caaggcacagg gaatgcccgg gcccgggtgg ggaaggatca cggattcaca 13440
ccagtggaac accttgctaa gtttgcataa cgcgcacattt tatttgctac aacgcacgccc 13500
agaggttgcc cgcagccgcg ccaccccggtt attagatttgc atcaagacag cgttgcgc 13560
ccatccacccg caaaaacagg cgtatgggtt gacattaccc acttcagtgc tgggtatcgc 13620
cgacacgat actaatctgg caaatctgg cggcgcactg gagtcact ggacgttcc 13680
cggtcagccg gataacacgc cgcagggtgg tgaactgggtg tttgaacgct ggcgtcggt 13740
aagcgataac agccagtgg ttcagggttc gctggcttc cagactttac agcagatgcg 13800
tgataaaacg cgcgtgtcat taaatacgcc gcccggagag gtgaaactgaa ccctggcagg 13860
atgtgaagag cgaaatgcgc agggcatgtg ttcgttggca ggtttacgc aaatcgtaa 13920
tgaagcacgc atacccgctt gcaagttgtt agtacccgg ggatcacaac ttgcctctg 13980
aagaggaaga acagaaggat gcccacaactc tccgtgtggc tactctccag tgggttcatc 14040
ttacttctga tggcatttcc ctctagaaag tgctactatc atccacacat ttctacatc 14100
gaccacccaa aggaccctcc caaattctct tccctctctga gtagtctcca caccgttac 14160
caccatccca gaattaaaaat cctaaactgca ctctggcggtg tgacttgcct cagtccttgc 14220
aataagagtt gttggcagtg ccaggcgtgg tggcgcacgc cttaattcc agcacttggg 14280
aggcagagggc aggccggatt ctgaggctga ggccagcctg gtctacagag tgagttccag 14340
gacagccagg gctatacaga gaaacccgtt gtcgaaaaac caaaaaaaaaaaaaaaagg 14400
gttggcagag tgggtttat ataccagggtt gagatttcaaa atgagttggct gaagctgttag 14460
ccagaaggaa cttagaggat agtcataaac ttaaaaagaa atgtagagag tagcagaaac 14520
attgagagag tgggcacaca gccactgtgt gatgtggca gacacatac cagccagcta 14580
tacatgcata agtgtatatt ggcccatcc tgactgtatc gacacaggaa aacagataga 14640
cgggtttagg tggccatggc ctcccttgc tgccttcc taagggtcat ctcaagacct 14700
tatgtctctt taacttccattt attgtactt agcttcttaga tatcacctcc agattgtct 14760
ccttgggtac atcgtgtatc ctgggtatcc ctagggcttc ctgatccat ctgtcata 14820
gaggctgcaaa ctaaagaggtt cttcttataa cttcacaccc tgcacccaa aggaagacac 14880
agaagttcac agaggtgaag tgatgtatc aggacatata gtcgacaaagc atcagggtcc 14940
ggattatctg actctactt aactttatgt taaatgtgtt ttatgtcattt aacactgtca 15000
ttcctgtgtt tcaatgtctgg gagactccca agcactctt ggcacaagcc acaattaagg 15060
gactctgaca ctctgcattt attaatttagc atgggtgtt ctatgttcc agattcatga 15120
ttgtttactt tccatataatg gtcgtatcagg gttgtggaa attttttggg gacagaattt 15180
gaggcaatcc acctcttcga ggaagccctt atctggaaaa gcttacaact cagggacagt 15240
aactgttaggc ccagtccttgc gttgtccaaa tgggtttat ggtttgaatc tgcaagcc 15300
tccatgtgtt cttttttttt aacatggacg ctccttgc taaactgttta tggaggc 15360
ttgagactgg atgtcttttgc tccatgtt ttgttacatc atctgtcaag atatgaccc 15420

ggcattgtac cagctaccac agactatgcc tctccagctt tcatgttctc cccaccatga 15480
 tagacttgtatctccataaa atggaatcaa agcaaaacttt tccctgcatta agtttttttt 15540
 tttctgttaa gtgttggtc acagggacaa gaaaacactc aatacagata attagtacca 15600
 gagttgaggt tcattgtct agcaagttgg atcaaatttt tagggctttg gaactgattt 15660
 ataagagaca tggtagaagag tctgaagctg tgggctacag aagtgtcacc agtttttaag 15720
 aatagtttaa tacaccatgg gaattgtgaa aatcagaatg ctcacacaaa ggcagacagg 15780
 aaaacgttag catgtggcgt gtgagagggc ataagaagga acctaggggg aaatgagcta 15840
 gaagccattc ggctacgtta gggAACGTGT gtggctgtgc ttggcccatg ccctggcaat 15900
 ctgaatgagg ccaaattttt aaggagtggaa ctaactcgat tgcagagaa aatatcaaga 15960
 cagaccacca ctcaggctat gccgtgtttg tgaccgacca gctactcttta gccagctcta 16020
 ttgtgaaatt ccagagcaat tacagagca tgaagataca tacagtttag tgaagtaagg 16080
 ggtgtgggtc cctaagtggaa tgggtgcataa atctatgttag gtgatgccta agtgacactt 16140
 gataatccaa aatatcagca atgtggaaatg tcttccaagg agacctgttag acacacattt 16200
 tagaacttttgc tctatggctg taataaatacg ttagctagaa atcatttcctt gaagaggtt 16260
 gtctgagttt cgggtccagg gcaaaacattt cgtatggca aggaaggcat tgcagtcagg 16320
 agccaaagggt cagctgtca cattgcatac agagtagaga gtcagagtgt gaggtagaaag 16380
 aggatacagg ttataaaacc tcaactgtcca ctctcagcaa tccattttctt cctaaaaggc 16440
 tttaccttctt aaagatttttgc tcttcaaaa ccagtaccag taggcctggaa acaaaagggt 16500
 aaacaaatatgaa gccttggcggcatttccac acttaaaaca gggcatcacc taggaggagc 16560
 cctgtgtgca gtaggaagtg tggcctctgt gtcaggaatg ctcaggctaa taaggggtcc 16620
 tctatctgag ggaccctatg aagattcaac aagttagttgt gagaattccc tggtaatgg 16680
 tgctaccaat ttgacatttg tagacctgctt attgtgtgct tcttattttgg gctctcccat 16740
 ctcccaactt tccaaacccat attcacatattt aatcccttcc accaccatgc aacacttagt 16800
 aggagagaag gaaggttaga agagaaaagtgg tttatgtatc tatttagact acttcctgct 16860
 gattaggggc aagtccaaatc gtcattgtca ggatacctcc aaccagcaac cagcaaaacca 16920
 gcaaaatcaga aacagcaaaa gcagccaaaca aggcagcact aaccagcagg attggggtcg 16980
 gtagcgtggg agcagtcact actggcttc tcatggctt ggcatataata ctctctcaag 17040
 aaattccgtat ttttttccc aaccacctgaa aattccgtaa ttttaatgc aaactatcta 17100
 cagctggcaa aaatcacatc tctcttagag cacaagacaa atcatagttt ctggcttattt 17160
 gcaatctgaa gcatctcaat atccccacacc tgggattaaa acaaaaacat attcacatca 17220
 cataactgtt tttttttcc aattttttat taggtatccc ttttatttac atttcaaatg 17280
 ctatcccgaa agtcccttat accctccac ctccctgctc ccctacacac ccactccac 17340
 ttttgaccc tggagttccc cggtaactggg gcatataaaag tttgcaagac caaggggcct 17400
 ctcttcccgag ttagtggccga ctaagccatc ttctgtcata tatgcagata gagacacag 17460
 ctctgggggtt acttagttgtt tcatattgtt gttccaccta tagggtcgca gacccttca 17520
 gtccttggg tactttgtct agtcctccca ctgggggctc tgggttttat ctaatagatg 17580
 actgtgagca tccacttctg tatttgacag gcaactggctt agcgtcacat gagccagcta 17640
 tattcagggtc ttttcagcaa aaccttgcgt gcatgtgcaaa tagtgcgtc gtttgggtgt 17700
 tgattatggg atggatccac tagttctaga gcccggcca cccgggtggaa gctccagctt 17760
 ttgttccctt tagtgagggt taatttgcgt gtttggctaa tcatggctat agctgttcc 17820
 tggtagaaat tggtagaaatccgc tcacaattcc acacaacata cgagccggaa gcataaaagt 17880
 taaaaggctgg ggtgctaat gaggtagcta actcacatatttgcgtt gctcaactgcc 17940
 cgctttcccgag tggggaaaacc tggtagccca gctgcattaa tggatggcc aacgcgcggg 18000
 gagaggcggtt tggcgatattt ggcgttcccgat cgttccctcg ctcaactgact cgctgcgtc 18060
 ggtcggttccgg ctgcggcgttccggat cgttccctcgat tcaactcaaaag gggtaataac gtttatccac 18120
 agaattcagggtt gataacgcgtt gaaagaacat gtgagcaaaa ggcagcaaa aggcaggaa 18180
 ccgtaaaaaaag ggcgttccggat tggcgatattt ccatacgctc cgccccctg acgagcatca 18240
 caaaaatcgaa cgctcaagtc agagggtggcg aaacccgaca ggactataaa gataccaggc 18300

gttccccctt ggaagctccc tctgtcgctc tccctgttccg accctgccgc ttaccggata 18360
 cctgtccgccc ttctccctt cgaaaaagcgt ggcgccttct catacgctcac gctgttaggtt 18420
 tctcagttcg gtgttaggtcg ttgcgtccaa gctgggctgt gtgcacgaac ccccccgttca 18480
 gccccgaccgc tgccgccttat ccggtaacta tctgtttagt tccaaacccgg taagacacga 18540
 cttatcgcca ctggcagcag ccactggtaa caggattagc agagcgaggt atgtaggcgg 18600
 tgctacagag ttcttgaagt ggtggcctaa ctacggctac actagaagga cagtatttgg 18660
 tatctgcgtc ctgctgaagc cagttacccctt cggaaaaaaga gttggtagct cttgatccgg 18720
 caaacaacc accgctggta gcgggtggttt ttttgttgc aaggcagcaga ttacgcgcag 18780
 aaaaaaaagga tctcaagaag atcccttgat cttttctacg gggctgacg ctcagtggaa 18840
 cgaaaaactca cgttaaggga ttttggtcat gagattatca aaaaggatct tcacctagat 18900
 cctttaaat taaaaatgaa gttttaaatc aatctaaagt atatatgagt aaacttggtc 18960
 tgacagttac caatgcttaa tcagtggatc acctatctca gcgtatctgtc tatttcgttc 19020
 atccatagtt gcgtgactcc ccgtcgtgtt gataactacg atacgggagg gcttaccatc 19080
 tggcccccagt gctgcaatga taccgcgaga cccacgctca ccggctccag atttacgc 19140
 aataaaaccag ccagccggaa gggccgagcg cagaagtggt cctgcaactt tatccgcctc 19200
 catccagtctt attaattgtt gccgggaagc tagagtaagt agttcgccag ttaatagttt 19260
 ggcgcacgtt gttgccattt ctacaggcat cgtgggtgtca cgctcgctgt ttggtatggc 19320
 ttcattcage tccgggtcccc aacgatcaag gcgagttaca tgatccccca tgggtgtgaa 19380
 aaaagcggtt agtccttcgt gtcctccgtat cgttgcgtt agtaagttgg ccgcagtgtt 19440
 atcactcatg gttatggcag cactgcataa ttctcttact gtcatgcccattt ccgtaaagatg 19500
 cttttctgtt actgggtgagt actcaacccaa gtcattctga gaatagtgtt tgccggcggacc 19560
 gagttgtctt tgccggcgtt caatacggga taataccgcg ccacatagca gaactttaaa 19620
 agtgcgtatc attggaaaac gttttcgggg gcgaaaaactc tcaaggatct taccgcgttt 19680
 gagatccagt tcgatgttac ccactcggtt acccaactga tcttcagcat cttttactttt 19740
 caccagcggtt tctgggtgag caaaaacagg aaggcaaaaat gcccggaaaa agggaataag 19800
 ggccgacacgg aaatgttggaa tactcataact cttccctttt caatattttt gaagcattta 19860
 tcagggttat tgcgttcatga gcccggatcat atttgaatgt atttagaaaa ataaacaaat 19920
 aggggttcccg cgcacatttc cccggaaaatg gcccacccaa ttgttaagcgt taatattttt 19980
 ttaaaaattcg cgttaaaaattt ttgttaaattt agtcattttt ttaaccaata gggccggaaatc 20040
 ggccaaaatcc cttataaaatc aaaagaatag accggatagat ggttggatgt tggccagtt 20100
 tggaaacaaga gtccactatt aaagaacgtg gactccaaacg tcaaaaggcg aaaaacgcgc 20160
 tatcaggggcg atggcccact acgtgaacca tcaccctaat caagttttt ggggtcgagg 20220
 tgccgtaaatg cactaaatcg gaaccctaaa gggagcccc gatttagagc ttgacggggaa 20280
 aagccggcga acgtggcgag aaaggaaggg aagaaagcga aaggagcggg cgctaggccg 20340
 ctggcaagtg tagcggtcac gtcgcgtat accaccacac ccggccgcgt taatgcggcc 20400
 ctacaggggcg cgtcccatcc gcatttcagg ctgcgtactt gttgggaagg gcgatcggtg 20460
 cggggccttcc cgttattacg ccagctggcg aaagggggat gtgtgtcaag gcgatataatg 20520
 tgggttaacgc cagggttttc ccagtcacga cgttgcgtttt ccacggccag tgagcgccg 20580
 taatacgact cactataggg cgaattgggtt accggggccccccc 20623

<210> 2
 <211> 6708
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: R15/APPA +

intron plasmid with pBLCAT3 vector

<400> 2
ggatccccctt tgctatgtag ttttaatgg aaattacaac ccatagtgtg ttgataaata 60
gagagtccctg tttggtttaa gcaacctctg tttctcataa actccataaa aacaggaata 120
ctctttgttt ctgcataac caaaagattt agtgaattga aaacaatgtt cccttagagt 180
ataggtctaa taaccccgaa aatattacca tgatactgag catttgcataa tatctcatag 240
catgtatccat ccatagtcca tcaatgagag agacattaa catgatttc attaatcagg 300
tggaaaagac atgacaacat tcacaggcac tgcacagaac atagtggtcc accttgcaca 360
tatttcacta aactaggttt atctattttg ttgcattctc taacatctct gcaatgaagc 420
aggtcaacag tgccacatat ccttactta acctaaggaa cacaaaaaat tttctacata 480
tattctgtt agagagtgt taaaataagt tttccaagaa tggaaaagaa atgttctgac 540
ttaacaatta agacagtatt tatttaaagc aagaatatg aggcacacaa gaaaatattt 600
tggaaagaaa ccatttgggt aacaatattt caaataaaaa tagacaaca tagttaattg 660
taaaacatat gtttgcacc cccttcttt caataggctt aatgtgaata aatgttaaa 720
gattctctt gggtggtgc aaattgtcca cgaataagac aaaatataaa aataaggact 780
gagtctcaca aaatgaaaag gaaatatatt cagaaagaga atcttgagag aatgtgtgt 840
cacaattaa agaaaacctg tggtaatga catctgagg cctgagctat tactgacatt 900
taagataaag gtaactgtat acattgtcc cattgagggg acaagaaagc tgctctcatg 960
ttcagctcta taattcttgc cttaaacaac ttaaatagaa tgatttaaaa tatggagctg 1020
tccatggacc tttgaaatataaaatagtc aagcaacttat caaggaatta cagattcatt 1080
gatactaaca caggtaaatac ccacacgtgt tttgagacta catttgcgtt gattttattt 1140
atgtaatagg tcacatgtt ttcggccaa tgggtgtt attcggttac ttcaagagaa 1200
tagtggcaac tgatgctatg tattctaggg gtttgaagtg atgtttcatg attgaaattt 1260
gtaaaagaat aacatcatca ttcttaacaa tagaacatataaagtcacac agaagtgaca 1320
gtgtttaagc tgtactattt atcaaagaaaa tttattacct tcagttcaaa tggaaataat 1380
tactgataat acaaacatgt gtgaacacac actaattctt tccaaatgca cagtgatata 1440
cagaaaatataat tagcaagtag aatgcaatataat tttatataacg attgtattt tcaatcaatt 1500
gtatgtatca atatatgggc tattttctt cacatgattt tattcaaaatt tactctaaatc 1560
attgttgaac catttagaaa aggcatactg gcaacttttc cttacctcat ccagctggc 1620
aaaagtccca gtgtggagta aaggatgcaa gatttctgc tctgttaagt ataaaataat 1680
agtatgaatt caaagggtgcc attcttctgc ttcttagttt aaaggcagtg cttgttctt 1740
ccagcacaga tctggatctc gaggagctg gcgagatttt caggagctaa ggaagctaaa 1800
agccgccacc atgaaaagcca tcttaatccc atttttatct cttctgatcc cgttaacccc 1860
gcaatctgca ttgcctcaga gtgagccgga gctgaagctg gaaagtgtgg tgattgtcag 1920
tcgtcatggt gtgcgtgctc caaccaaggc cacgcaactg atgcaggatg tcaccccaaga 1980
cgcatggcca acctggccgg taaaactggg ttggctgaca cgcgcgggtg gtgagctaat 2040
cgccatctc ggacattacc aacgccagcg tctggtagcc gacggattgc tggcgaaaaa 2100
gggctgcccc cagtcggc tattgtgtt gtcgacgagc gtacccgtaa 2160
aacaggcgaa gccttcgccc cggggctggc acctgactgt gcaataaccg tacataacca 2220
ggcagatacg tccagtcggc atccgttatt taatctctt aaaaactggcg tttgccaact 2280
ggataacgca aacgtgactg acgcgatctt cagcaggcggca ggagggtcaa ttgctgactt 2340
taccgggcat cggcaaacgg cggttcgca actggaaacgg gtgcttaattt tccgcataatc 2400
aaacttgc tttttttttt agaaacaggaa cgaaagctgt ttcattaaacgc aggcattacc 2460
atcgaaactc aagggtgactg ccgacaatgt ctcattaaacc ggtgcggtaa gcctcgatc 2520
aatgctgacg gagatatttc tcctgcaaca agcacaggaa atgcccggagc cgggggtgggg 2580
aaggatcacc gattcacacc agtggaaacac cttgcttaagt ttgcataacg cgcaattttt 2640
tttgcataaa cgcaacggcag aggttgcggc cagccgcggcc acccccgttat tagatttgc 2700

tcgtcgttg gtatggcttc attcagctcc ggttccaaac gatcaaggcg agttacatga 5640
 tccccatgt tgcggaaaa agcggttagc tccttcggtc ctccgatcgt tgcagaagt 5700
 aagttggccg cagtgttac actcatggt atggcagcac tgcatataattc tcttactgtc 5760
 atgccatccg taagatgctt ttctgtgact ggtgagttact caaccaagtc attctgagaa 5820
 tagtgtatgc ggccgaccgag ttgccttgc cccgcgtcaa tacgggataa taccgcgcca 5880
 catagcagaa cttaaaaagt gctcatcatt ggaaacgtt cttcgccccg aaaaactctca 5940
 agatcttac cgctgtttag atccagttcg atgtaaacca ctcgtgcacc caactgatct 6000
 tcagcatctt ttactttcac cagcgtttctt gggtgagcaaa acacaggaag gcaaaatgcc 6060
 gcaaaaaagg gaataaggc gacacggaaa tggtaatac tcatactctt ctttttcaa 6120
 tattattgaa gcatttatca gggttattgt ctcatgagcg gatacatatt tgaatgtatt 6180
 tagaaaaata aacaaatagg ggttccgcgc acatccccca gaaaagtgcc acctgacgtc 6240
 taagaaacca ttattatcat gacattaacc tataaaaata ggcgtatcac gaggcccttt 6300
 cgtctcgccg gttcggta tgacggtaa aacctctgac acatgcagct cccggagacg 6360
 gtcacagctt gtctgttaagc ggtgtccggg agcagacaag cccgtcaggg cgcgtcagcg 6420
 ggtgttggcg ggtgtccggg ctggcttaac tatgcggcat cagagcagat tgcgttgc 6480
 gtgcaccata tgccgtgtga aataccgcac agatgcgtaa ggagaaaata ccgcattcagg 6540
 cgcattcgc cattcaggct ggcgttgc acatgcgttgc ggcgttgc 6600
 ctattacgcc agctggcgaa agggggatgt gctgcaaggc gattaagttt ggttaacgcca 6660
 gggtttccc agtcacgacg ttgtaaaacg acggccagtg ccaagctt 6708

<210> 3
 <211> 4060
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: R15/APPA +
 intron transgene

<400> 3
 ggatccccctt tgctatgttag ttttaatgg aaattacaac ccatagtgtg ttgataaata 60
 gagagtccctg tttggtttaa gcaacctctg tttctataa actccataaa aacaggaata 120
 ctctttgtt ctagcataac caaaagattt agtgaattga aaacaatgtt cccttagagt 180
 ataggcttaa taaccccgaa aatattacca tgatactgag catttgcgttatctcatag 240
 catgtgtat ccatagtcca tcaatgagag agacattaa catgattttc attaatcagg 300
 tggaaaagac atgacaacat tcacaggcac tgcacagaac atagtggcc accttgacaca 360
 tatttcacta aacttaggtt atctattttgc ttgccttctc taacatctct gcaatgaagc 420
 aggtcaacag tgccacat ccttactta acctaaggaa cacaaaaaat tttctacata 480
 ttcctgggtt agagagtgt taaaataagt tttccaagaa tggaaaagaa atgttctgac 540
 ttaacaattt agacagtatt tatttaaagc aagaaatatg aggcacacaa gaaaatattt 600
 tggaaagaaa ccatttggtg aacaatattt caaataaaaa tagacaaaca tagttattt 660
 taaaacatat gtttggccag cccttctttt caataggctt aatgtgaata aaatgttaaa 720
 gattctctt ggggtggctgc aaattgtcca cgaataagac aaaatataaa aataaggact 780
 gagtctcaca aatgaaaag gaaatattt cagaaagaga atcttggagag aatgtgttgc 840
 cacaattaa agaaaacctg tggtaatga catcctgagg cctgagctat tactgacatt 900
 taagataaag gtaactgtat acatttgc cattggggg acaagaaaagc tgctctcatg 960
 ttcagctcta taattttgc cttaaacaac taaaatagaa tgattttaaa tatggagctg 1020

tccatggacc ttgaaatataaaatgtca agcaacttat caaggaatta cagattcctt 1080
 gatactaaca caggtaaatac ccacacgtgt tttgagacta catttgcgtt gattttattt 1140
 atgtaatagg tcacatgttt ttcggccaa tggtgcgtt attcggttac ttcaagagaa 1200
 tagtggcaac tgatgcgtat tattctaggg gtttgaagtg atgtttcatg attgaaatattt 1260
 gtaaaagaat aacatcatca ttcttaacaa tagaacatataaagtcacac agaagtgaca 1320
 gtgtttaagc tgtactattt atcaaagaaaa tttattacct tcaagttcaa tggaaataat 1380
 tactgataat acaaacatgt gtgaacacac actaatccta tccaaatgca cagtgataca 1440
 cagaaaatata tagcaagtag aatgcaatataat tataatataacg attgtatttt tcaatcaatt 1500
 gtatgtatca atatataggc tattttctta cacatgattt tattcaaaatt tactctaattc 1560
 attgttgaac catttagaaa aggcatactg gcaacttttcc tttacctcat ccagctggc 1620
 aaaagtccca gtgtggagta aaggatgcaaa gatttccgc tctgttaagt ataaaataat 1680
 agtatgaatt caaagggtgcc attcttctgc ttcttagttt aaaggcagtg cttgcttctt 1740
 ccagcacaga tctggatctc gaggagctt gcgagatttt caggagctaa ggaagctaaa 1800
 agccgcccacc atgaaagcca tcttaatccc attttatctt cttctgattt cgttaacccc 1860
 gcaatctgca ttcgctcaga gtgagccgga gctgaagctg gaaagtgtgg tgattgttag 1920
 tctgtatggt gtgcgtgctc caaccaaggc caccaactg atgcaggatg tcacccaga 1980
 cgcacatggcca acctggccgg taaaactggg ttggctgaca cccgcgggtt gtgagctaat 2040
 cgcctatctc ggacattacc aacgcccacg tctgttagcc gacggatttgc tggcgaaaaa 2100
 gggctgcccc cagtcgttgc aggtcgcgtat tattgtgtat gtcgacgagc gtacccgtaa 2160
 aacaggcgaa gccttcggc cccggctggc acctgactgt gcaataaccc tacataacca 2220
 ggcagatacgt tccagtcccg atccgttattt taatcctcta aaaaactggcg tttgccaact 2280
 ggataacgcg aacgtgactg acgcgtatctt cagcaggcga ggagggtcaa ttgctgactt 2340
 taccgggcat cggcaaacgg cgtttcgcga actgaaacgg gtgcttaattt ttccgcaatc 2400
 aaacttgc ttaaaacgtt agaaacagga cggaaactgt tcattaacgc aggcatattt 2460
 atcggaaactc aaggtagcgccg ccgacaatgtt ctcatttacc ggtgcggtaa gcctcgatc 2520
 aatgtgtacg gagatatttcc ttctgcaaca agcacaggga atgcggagc cgggggtgggg 2580
 aaggatcacc gattcacacc agtggaaacac cttgcttaagt ttgcataacg cgcaattttt 2640
 tttgctacaa cgcacgcccag aggttgcggc cagccgcgc accccgttat tagatttgat 2700
 caagacacgc ttgacgcccc atccacccgca aaaaacaggcg tatgggtgtga cattacccac 2760
 ttcagtgcgtt tttatgcggc gacacgatac taatctggca aatctggcg ggcacactgga 2820
 gctcaactgg acgttcccg gtcagccggtaa acacacgc cccgggttgc aactgggttt 2880
 tgaacgctgg cgtcggtctaa ggcataacag ccagtggattt caggtttcgc tggcttcca 2940
 gactttacag cagatgcgtg ataaaacgc gctgtcatta aatacgcgc cccggagaggt 3000
 gaaactgacc ctggcaggat gtgaagagcg aatgcgcag ggcacgtgtt cgttggcagg 3060
 ttttacgcaaa atcgtgaatg aagcacgcac acccgcttgc agtttgcataatgttgc 3120
 gttattgggtt cccttaaacgc cctgggtctta cgcctgaaata agtgcataatgttgc 3180
 tggcagaaat tgcggatc tttgtgaagg aaccttactt ctgtgggtgtt acataattgg 3240
 acaaactacc tacagagatt taaagctcta agtgcataatgttgc aatgcgcac 3300
 gtgttaaact actgattctt atttttgtt tattttgcataatgttgc aatgcgcac 3360
 aatgggagca gtgggtggat gccttaatgttgc aatgcgcac 3420
 catcttagtgc tgcgtggatc actgctgttgc tcaacatcc tactccttca aaaaagaaga 3480
 gaaaggtaga agaccccaag gactttcctt cagaattgtt aagtttttttgc aatgcgcac 3540
 tgtttagtaa tagaactctt gcttgcctt ctatttacac cacaaaggaa aatgcgcac 3600
 tgctataacaa gaaaattatgtt gaaaatatttcttgcataacccat tataagttagg ctttttttgc 3660
 ataatcataa catactgttt tttcttactc cacacaggca tagagtgttgc ttttttttgc 3720
 actatgtca aaaattgtgtt accttttagt ttttaatttgc taaagggtt aataaggaat 3780
 atttgcgtt tagtgccttgc actagagatc ataatcgcac ataccacatttgc ttttttttgc 3840
 ttacttgcctt taaaaaacctt cccacacccctt cccctgaaacc tggaaacataa aatgcgcac 3900

attgttgtt ttaacctgtt tattgcagct tataatggtt acaaataaag caatagcatc 3960
acaaatttca caaataaagc attttttca ctgcattcta gttgtggttt gtccaaactc 4020
atcaatqtat cttatcatqt ctqqatcqat cccqqqgtac 4060

<210> 4
<211> 6116
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: R15/APPAl
plasmid with pBLCAT3 vector

cgcacatggcca acctggccgg taaaactggg ttggctgaca ccgcgcggtg gtgagctaat 2040
 cgcctatctc ggacattacc aacgcgcagcg tctggtagcc gacggattgc tggcgaaaaa 2100
 gggctgccc cagtctggtc aggtcgcgat tattgctgat gtcgcacgagc gtacccgtaa 2160
 aacaggcgaa gccttcgccc cccggctggc acctgactgt gcaataaccg tacataacca 2220
 ggcagataacg tccagttcccg atccgttatt taatcctcta aaaactggcg tttgccaact 2280
 ggataacgcg aacgtgactg acgcgcattc cagcaggca ggagggtcaa ttgctgactt 2340
 taccgggcat cggcaaacgg cgttgcgca actggAACgg gtgcctaatt ttccgcatac 2400
 aaacttgtgc cttaaacgtg agaaacagga cgaaagctgt tcattaaacgc aggcattacc 2460
 atcggaactc aaggtgagcg ccgacaatgt ctcattaaacc ggtgcggtaa gcctcgcatc 2520
 aatgctgacg gagatatttc tccgtcaaca agcacaggga atgcccggagc cgggggtgggg 2580
 aaggatcacc gattcacacc agtggAACac cttgcttaagt ttgcataacg cgcaatttt 2640
 tttgctacaa cgcaacgcag agttgccc cagccgcgc accccgttat tagatttgat 2700
 caagacagcg ttgacgcccc atccaccgc aaaaacaggcg tatggtgtga cattaccac 2760
 ttcagtgctg tttatcgccc gacacgatac taatctggca aatctcgccg ggcactgga 2820
 gctcaactgg acgctcccg gtcagccggta aacacgcgc ccaggtggtg aactgggtt 2880
 tgaacgctgg cgtcgctaa gcgataacag ccagtggatt caggttgc tggctttcca 2940
 gactttacag cagatcgctg ataaaaacgcgc gctgtcatta aatacgcgc ccggagaggt 3000
 gaaactgacc ctggcaggat gtgaagagcg aaatgcgcag ggcattgtt cgttggcagg 3060
 ttttacgcaa atcgtaatg aagcacgcac acccgcttgc agttttaag gtataaggca 3120
 gttattggtg cccttaaacg cctggtgcta cgccctgata agtgataata agcggatgaa 3180
 tggcagaaat tgcggcgtc tttgtgaagg aaccttactt ctgtgggttg acataattgg 3240
 acaaactacc tacagagatt taaaaaacct cccacaccc cccctgaacc taaaacataa 3300
 aatgaatgca attgtgttg ttaacttggt tattgcagct tataatgggtt acaaataaag 3360
 caatagcattc acaaatttca caaataaaggc attttttca ctgcatttca gttgtggttt 3420
 gtccaaactc atcaatgtat cttatcatgt ctggatcgat ccccggtac cgagctcgaa 3480
 ttcgtaatca tggcatagc tttccctgt gtgaaattgt tattcgctca caattccaca 3540
 caacatacga gccgaaagca taaaatgttaa agcctgggtt gcttaatggag tgagctaaact 3600
 cacattaatt gcttgcgtc cactgcccgc tttccagtcg ggaaacctgt cgtgccagct 3660
 gcattaaatgaa atcgccaaac ggcgggggag aggccgtttt cgtattggc gctttccgc 3720
 ttcctcgctc actgactcg tgcgctcggt cgttccggc cggcgagcgg tatcagctca 3780
 ctcaaaggcg gtaatacggt tatccacaga atcaggggat aacgcaggaa agaacatgt 3840
 agcaaaaaggc cagcaaaagg ccaggaaccc taaaaggcc gcttgcgtgg cgaaaaatcca 3900
 taggctccgc cccctgacg agcatcaca aatcgacgc tcaagtgcaga ggtggcgaaa 3960
 cccgacagga ctataaaagat accaggcggt tccccctggaa agtccctcg tgcgctctcc 4020
 tggccgacc ctgcccgtt ccggataaccc tgcgccttt ctcccttgcg gaagcgtggc 4080
 gtttctcaa tgctcaacgt gtaggtatct cagttccggc tagtgcgttgc gctccaagct 4140
 gggctgtgtg cacaaccccc ccgttcagcc cgaccgctgc gccttateccg gtaactatcg 4200
 tctttagtcc aacccggtaa gacacgactt atcgccactg gcacgcggca ctggtaacag 4260
 gattagcaga gcgaggatg taggcgggtc tacagatgc ttgaagtggt ggcctaacta 4320
 cggctacact agaaggacag tattttggat ctgcgcctcg ctgaagccag ttacccctgg 4380
 aaaaagaggat ggtagctttt gatccggca acaaaccacc gctggtagcg gtggttttt 4440
 tggccgttcaag cagcagatta cgcgcagaaa aaaaggatct caagaagatc ctttgcattt 4500
 ttctacgggg tctgacgctc agtggAACga aaactcactt taaggattt tggcatgag 4560
 attatcaaaa aggatcttca cctagatccct tttaaattaa aatgaagtt taaaatcaat 4620
 ctaaaatgtata tattttttttt cttgggtctga cagttacca tgcttaatca gtgaggcacc 4680
 tatctcagcg atctgtctat ttcgttcatc catagttgcg tgactcccg tcgtgttagat 4740
 aactacgata cgggagggtt taccatctgg ccccgatgtc gcaatgatac cgcgcagaccc 4800
 acgctcaccg gctccagattt tattttttttt aaaccacca gccggaaaggc cgcgcgcag 4860

aagtggtcct gcaactttat ccgcctccat ccagtcatt aattgttgcc gggaaagctag 4920
 agtaagtagt tcgcccagtta atagttgcg caacgttgcg gccattgcta caggcatcgt 4980
 ggtgtcacgc tcgtcggttgcg gttcccaac gatcaaggcg 5040
 agttacatga tccccatgt tgcggaaac agcggttgcg tccctcggtc ctccgatcgt 5100
 tgcagaagt aagttggccg cagtgttgcg actcatgtt atggcagcac tgcataattc 5160
 tcttactgtc atgcacatccg taagatgctt ttctgtgact ggtgagact caaccaagtc 5220
 attctgagaa tagtgtatgc ggcgaccgag ttgcgttgc cccgcgtcaa tacggataa 5280
 taccgcgcca catagcagaa cttaaaaagt gctcatcatt ggaaaacgtt ctgcggccg 5340
 aaaactctca aggcattac cgctgttgcg atccagttcg atgtacccca ctgcgtgcacc 5400
 caactgatct tcagcatctt ttactttcac cagcggttgc ggtgagcaaa acacaggaa 5460
 gcaaaatgcc gcaaaaaagg gaataaggc gacacggaa tggtaatac tcatactt 5520
 ccttttcaa tattattgaa gcatttatca gggttattgt ctcattgagcg gatacatatt 5580
 tgaatgttatt tagaaaaata aacaaatagg gttccgcgc acatccccca gaaaagtgcc 5640
 acctgacgtc taagaaacca ttattatcat gacattaacc tataaaaata ggcgtatcac 5700
 gaggcccttt cgtctcgccg gtttcggta tgacggtaa aacctctgac acatgcagct 5760
 cccggagacg gtcacagctt gtctgttgcg ggttccggg agcagacaag cccgtcagg 5820
 cgcgtcagcg ggtgttggcg ggtgttgggg ctggcttaac tatgcggcat cagagcagat 5880
 tgcgttgcg gtcacccata tgcgttgcg aataccgcac agatgcgtaa ggagaaaata 5940
 ccgcatttcg cgcatttcg cattcaggct ggcactgt tggaaaggc gatcggtgcg 6000
 ggcctttcg ctattacgaa agctggcgaa agggggatgt gctgcaaggc gattaagtt 6060
 ggttaacgcg gggtttccc agtcacgcg ttgtaaaacg acggccagtg ccaagc 6116

<210> 5
 <211> 3470
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: R15/APPA
 transgene

<400> 5
 ggatccccctt tgctatgttag ttttaatgg aaattacaac ccatagtgtg ttgataaata 60
 gagagtccctg tttggtttaa gcaacctctg tttctatcaa actccataaa aacaggaata 120
 ctctttgtt ctagcataac caaaagattt agtgaattga aaacaatgtt cccttagagt 180
 ataggtctaa taaccccgaa aatattacca tgatactgag catttgcgtaa tatctcatag 240
 catgtatccat ccatagttcca tcaatgagag agacattaa catgatccc attaatcagg 300
 tggaaaagac atgacaacat tcacaggc ac tgcacagaac atagtggcc accttgcaca 360
 tatttcacta aactaggttt atctatccc ttgccttctc taacatctct gcaatgaagc 420
 aggtcaacag tgccacatcat ctttactta acctaaggaa cacaaaaat tttctacata 480
 tattcgtt agagaatgt taaaataatgtt tttccaaagaa tgaaaaagaa atgttctgac 540
 ttaacaatata agacagtatt tattaaagc aagaaatatg aggcacacaa gaaaatattt 600
 tggaaagaaa ccatttggtg aacaatattt caaataaaaa tagacaaaca tagttaattt 660
 taaaacatcat tttgaccag cccttctttt caataggctt aatgtgataa aatgttaaa 720
 gattctctttt gggtggctgc aaattgtccca cgaataagac aaaatataaa aataaggact 780
 gagtcctcaca aaatgaaaag gaaatataattt cagaaagaga atcttgagag aatgtttgtt 840
 cacaaatataa agaaaacctg tggtaatga catcctgagg cctgagctat tactgacatt 900

taagataaaag gtaactgtat acatttgc cattgagggg acaagaaagc tgctctatg 960
 ttcagctcta taattcttgc cttaaacaac ttaaatagaa tgataaaaa tatggagctg 1020
 tccatggacc tttgaaatat aaaatagtca agcaacttat caaggaatta cagattcctt 1080
 gatactaaca caggtaaatc ccacacgtgt tttgagacta catttgcgtt gatTTTATTG 1140
 atgtaatagg tcacatgttt ttcggccaa tggcgctgtt attcggttac ttcaagagaa 1200
 tagtgcaac tggatgtatg tattcttaggg gtttgaagtg atgtttcatg attgaaattt 1260
 gtaaaaagaat aacatcatca ttcttaacaa tagaacatataa aagtcacac agaagtgaca 1320
 gtgtttaaagc tggatgtatg atcaaagaaa ttttattacct tcagtttcaa tggaaataat 1380
 tactgataat acaaacatgt gtgaacacac actaattccta tccaaatgca cagtgatata 1440
 cagaaaatat tagcaagtag aatgcaataat ttatataacg attgtatttta tcaatcaatt 1500
 gtatgtatca atatatgggc tattttctta cacatgattt tattcaatt tactctaattc 1560
 attgttgaac catttagaaa aggcatactg gcaacttttc cttacctcat ccagctggc 1620
 aaaagtccca gtgtggagta aaggatgcaa gatttgcgtc tctgttaagt ataaaataat 1680
 agtatgaatt caaagggtgcc attttctgc ttcttagttt aaaggcagtg cttgtttttt 1740
 ccagcacaga tctggatctc gaggagctt gcgagatTTT caggagctaa ggaagctaaa 1800
 agccgccacc atgaaagcca tcttaatccc atttttatct cttctgattc cgtaaacc 1860
 gcaatctgca ttgcgtcaga gtgagccggg gctgaagctg gaaagtgtgg tgattgtcag 1920
 tcgtcatgtt gtgcgtgctc caaccaaggc cacgcaactg atgcaggatg tcaccccaga 1980
 cgcatggcca acctggccgg taaaactggg ttggctgaca cccgcgcgtg gtgagctaat 2040
 cgccatctc ggacattacc aacgcccggc tctggtagcc gacggattgc tggcgaaaaa 2100
 gggctgcccgg cagtcgtggc aggtcgcgtat tattgtgtat gtcgacgagc gtacccgtaa 2160
 aacaggcgaa gccttcgcgg cccggctggc acctgactgt gcaataaccg tacataccca 2220
 ggcagatacg tccagtcggc atccgttatt taatctcta aaaactggcg tttgccaact 2280
 ggataacgca aacgtgactg acgcgatcc cagcaggcggc ggagggtcaa ttgctgactt 2340
 taccggccat cggcaaacgg cggttcgcga actggAACGG gtgtttaatt ttccgcaatc 2400
 aaacttgcgtc cttaaacgtg agaaacagga cggaaagctgt tcattaaacgc aggcattacc 2460
 atcggaaactc aagggtgagcg cgcacaatgt ctcattaaacc ggtgcggtaa gcctcgcatc 2520
 aatgctgacg gagatatttc tcctgcaaca agcacagggg atgcccggagc cgggggtgggg 2580
 aaggatcacc gattcacacc agtggAACAC cttgcttaagt ttgcataacg cgcaattttta 2640
 tttgttacaa cgcaccccg aggttgcggc cagccgcgtt accccgttat tagatttgc 2700
 caagacagcg ttgacgcccc atccaccgc aaaaacaggcg tatgggtgtga cattaccac 2760
 ttcagtcgtc tttatcgccgg gacacgatac taatctggca aatctcgccg ggcactgg 2820
 gtcacttgg acgcttcccg gtcagccggta taacacccgg ccagggtggc aactgggttt 2880
 tgaacgcgtt cgtcggctaa gcgataacag ccagtggatt cagggttcgc tggctttcca 2940
 gactttacag cagatgcgtg ataaaacgc gctgttattta aatacgcgc cccggagaggt 3000
 gaaactgacc ctggcaggat gtgaagagcg aaatgcgcag ggcattgtgtt cggtggcagg 3060
 ttttacgcaa atcgtgaatg aagcacgcac acccgcttgc agtttgcataag gtataaggca 3120
 gttattgggtt cccttaaacg cctgggtgtca cgcctgttata agtgcataata agcggatgaa 3180
 tggcagaaat tcgcggatc tttgtgttgc aaccttactt ctgtgggtgtt acataattgg 3240
 acaaactacc tacagagatt taaaaaaaccc cccacaccc cccctgttacc tgaaacataa 3300
 aatgaatgca attgttgcgtt ttaacttgc tattgcagct tataatgggtt acaaataaaag 3360
 caatagcatc acaaatttca caaataaaagc attttttca ctgcatttca gttgtgggtt 3420
 gtccaaactc atcaatgtat ctatcatgt ctggatcgat ccccggttac 3470

<210> 6

<211> 5421

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: SV40/APPA +
intron plasmid

<400> 6

cgagatttc aggagctaag gaagctaaaa gcccacca tgaaagccat cttaatccc 60
tttttatctc ttctgattcc gttAACCCG caatctgcat tcgctcagag tgagccggag 120
ctgaagctgg aaagtgtggt gattgtcagt cgtcatggtg tgcgtctcc aaccaaggcc 180
acgcaactga tgcaggatgt caccccagac gcatggccaa cctggccggt aaaaactgggt 240
tggctgacac cgcgnggtgg ttagctaattc gcctatctcg gacattacca acgcccagcgt 300
ctggtagccg acggattgct ggcggaaaag ggctgcccgc agtctggtaa ggtcggcatt 360
attgctgatg tgcacgagcg taccctgtaaa acaggcgaag ccttcgcgc cgggctggca 420
cctgactgtg caataaccgt acataaccag gcagataacgt ccagtcccgta tccgttattt 480
aatcctctaa aaactggcgt ttgccaactg gataacgcga acgtgactga cgcgatcctc 540
agcaggcag gagggtcaat tgctgacttt accggcattc ggcaaaccgc gtttcgcgaa 600
ctggAACGGG tgcttaattt tccgcaatca aacttgcgc ttaaacgtga gaaacaggac 660
gaaagctgtt cattaacgc ggcattacca tccgaaactca aggtgagcgc cgacaatgtc 720
tcattaaccg gtgcggtaag cctcgcatca atgctgacgg agatatttct cctgcaacaa 780
gcacaggggaa tgccggagcc ggggtgggaa aggtcaccg attcacacca gtggAACACC 840
ttgctaagtt tgcataacgc gcaattttat ttgctacaac gcacgcccaga ggttgcgc 900
agccgcgcca ccccgattt agatttgatc aagacagcgt tgacgccccca ccacccgaaaa 960
aacaggcgtt tgggtgtgaca ttaccactt cagtgtgtt tattcgcgcgaa cacgataacta 1020
atctggcaaa tctcgccgc gcaactggac tcaactggac gcttcccggt cagccggata 1080
acacgcgcgaa aggtgtgaa ctgggtttt aacgctggcg tccgctaaac gataacagcc 1140
agtggattca ggtttcgctg gtcttccaga ctttacagca gatgcgtgat aaaaacgcgc 1200
tgtcattaaa tacgcgcgcg ggagaggtga aactgaccct ggcaggatgt gaagagcga 1260
atgcgcaggg catgtgttgc ttggcagggtt ttacgcaaat cgtgaatgaa gcacgcatac 1320
ccgcttgcag ttgttaaggc agttattggc gcccattaaac gcctggcgtt acgcctgat 1380
aagtgataat aagcgatgtc atggcagaaaa ttgcggat cttgtgaag gaaccttact 1440
tctgtgtgt gacataattt gacaaactac ctacagagat ttaaagctt aaggtaata 1500
taaaatttttt aagtgtataa tggtaaaac tactgattt aattgtttgt gtattttata 1560
ttccaaccta tggaaactgtatc gaatgggagc agtggggaa tgccttaat gaggaaaacc 1620
tgtttgctc agaagaaatg ccatctagtg atgatgaggc tactgctgac tctcaacatt 1680
ctactcctcc aaaaaagaag agaaaggtag aagacccaa ggactttct tcagaattgc 1740
taagttttttt gagtcatgtc gtgttttagta atagaactct tgcttgcctt gctatttaca 1800
ccacaaaggaa aaaagctgc ctgctataca agaaaattat ggaaaaatat tctgtaaacct 1860
ttataagtag gcataacagt tataatcata acatactgtt ttttcttact ccacacaggc 1920
atagagtgtc tgcttataat aactatgtc aaaaattgtg taccttttagc tttttttttt 1980
gtaaaagggtt taataaggaa tatttgcgtt atagtgccctt gactagagat cataatcagc 2040
cataccacat ttgttagaggt ttacttgc ttttttttttccacacactt cccctgttac 2100
ctgaaacata aaatgaatgc aattgttgc gttacttgc ttattgcgc ttataatgggt 2160
tacaaataaa gcaatagcat cacaatttc acaaataaag catttttttccactt 2220
agttgtgggtt tgcctaaactt catcaatgtc tcttgcgtt tctggatgc tcccccggta 2280
ccgagctcga attcgtaatc atggcatacg ctgtttctg tgcgttgcgc ttatccgc 2340
acaattccac acaacatacg agccggaaagc ataaagtgtt aagcctgggg tgcctaaatgt 2400
gtgagctaac tcacattaaat tgcgttgcgc tcactgccccg ctttccactc gggaaacctg 2460

tcgtgccagc tgcattaatg aatcgccaa cgcgccggga gaggcggtt gcgtattggg 2520
 cgctttccg cttccctcgct cactgactcg ctgcgctcg tcgttcggct gcggcgagcg 2580
 gtatcagctc actcaaaggc ggtataacag ttatccacag aatcagggggtaacgcagga 2640
 aagaacatgt gagcaaaaagg ccagcaaaag gccaggaacc gtaaaaaggc cgcttgcgt 2700
 gcgttttcc ataggctccg cccccctgac gagcatcaca aaaatcgacg ctcaagtcag 2760
 aggtggcgaa acccgacagg actataaaga taccaggcgt ttccccctgg aagctccctc 2820
 gtgcgctctc ctgttccgac cctgcccgtt accggatacc tgcgttgcgtt tctccctcgt 2880
 ggaagcgtgg cgctttctca atgctcacgc tgcgttgcgtt tcagttcggt gtaggtcggt 2940
 cgctccaagc tgggtgtgt gcacgaaccc cccgttcagc ccgaccgctg cgccttatcc 3000
 gttaactatc gtcttgcgtc caaccggta agacacact tatcgccact ggcagcagcc 3060
 actggtaaca ggattagcag agcgaggtat gtaggcgtt ctacagagtt cttgaagtgg 3120
 tggcttaact acggctacac tagaaggaca gtatggta tctgcgtct gctgaagcca 3180
 gttaccttcg gaaaaagagt tggttagctt tgatccggca aacaaccac cgctggtagc 3240
 ggtggttttt ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat 3300
 cctttatct tttctacggg gtctgcgtc cagtgaaacg aaaaactcact ttaaggatt 3360
 ttggtcatga gattatcaaa aaggatctt acctagatcc tttaaatta aaaaatgaagt 3420
 tttaaatcaa tctaaatgtt atatgagtaa acttggctc acagttacca atgcttaatc 3480
 agtgaggcac ctatctcagc gatctgtctt tttcgatcat ccatagttgc ctgactcccc 3540
 gtcgtgtaga taactacgt acggggggc ttaccatctg gcccagtgc tgcaatgata 3600
 ccgcgagacc cacgttcacc ggctccagat ttatcagcaa taaaccagcc agccggagg 3660
 gccgagcgcga gaagtggtcc tgcacttta tccgcctcca tccagtctat taattgtgc 3720
 cgggaagcta gagaatgtt ttcgcccgtt aatagttgc gcaacgttgc tgccattgtc 3780
 acaggcatcg tgggtgcacg ctgcgttgcgtt ggtatgtt cattcagctc cggtcccaa 3840
 cgatcaaggc gagttacatg atccccatg ttgtcaaaaa aagcggttag ctccctcggt 3900
 cctccatcg ttgtcagaag taagttggcc gcagtgtt cactcatgt tatggcagca 3960
 ctgcataatt ctcttactgt catgccatcc gtaagatgtc ttctgtgc tggtgagtag 4020
 tcaaccaagt cattctgaga atagtgtatg cggcgaccga gttgtcttgc cccggcgta 4080
 atacgggata ataccgcgc acatagcaga actttaaaag tgctcatcat tggaaaacgt 4140
 tcttcggggc gaaaactctc aaggatctt ccgctgttga gatccagttc gatgtaaacc 4200
 actcgtgcac ccaactgatc ttccatcttca ccagcggttgc tgggtgagca 4260
 aaaacaggaa ggcaaaatgc cgcaaaaaag ggaataaggg cgacacggaa atgttgaata 4320
 ctcataactct tccttttca atattattga agcatttac aggttatttgc tctcatgagc 4380
 gatacatat ttgaatgtat ttggaaaaat aaacaaatag gggcccgccg cacatttcc 4440
 cgaaaaatgc cacgtacgt ctaagaaacc attattatca tgacattaac ctataaaaat 4500
 aggcgtatca cgaggccctt tcgtctcgcc cggttcgggt atgacgggtga aaacctctga 4560
 cacatgcage tcccgagac ggtcacagct tgcgttgcgtt cggatgcggg gagcagacaa 4620
 gcccgtcagg ggcgtcagg ggggtttggc ggggtgtccggc gctggcttac ctatgcggca 4680
 tcagagcaga ttgtactgag agtgcacccat atgcgttgcgtt aaataccgca cagatgcgt 4740
 aggagaaaaat accgcatcg ggcgcattcg ccattcaggc tgcgcactgt ttggggagg 4800
 cgatcggtgc gggcccttgc gctattacgc cagctggcga aagggggatg tgctgcagg 4860
 cgatataagg tggtaacgc accgttttcc cagtcacgc gttgtaaaac gacggccagt 4920
 gccaagcttt acactttatc ttccggctc gtatgttgcgtt tggaaattgtg agcggataac 4980
 aatttcacac aggaacacgc tatgaccatg attacgaatt cggcgacgca ccatggctg 5040
 aaataacctc tgaaagagga acttgggttag gtaccttgcgtt aggccggaaag aaccagctgt 5100
 ggaatgtgtg tgcgttgcgtt tggaaatgttcccgatcccccggc tcccccagcag 5160
 aagcatgca tctcaatttgc tgcgttgcgtt tggaaatgttcccgatcccccggc tcccccagcag 5220
 gcaagaatgttcccgatcccccggc tcccccagcag 5280
 cgcccatccc gcccctaaactt ccggccatttcc tccggccatcc ggcgtactaa 5340

tttttttat ttatgcagag gccgaggccg cctcgccctc tgagctattc cagaagttagt 5400
gaggaggctc gaggagctt g 5421

<210> 7
<211> 17732
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Lama2/APPA
transgene

<400> 7
tcgagagtat ctttgcagc tgcgcctcca acaaagggtt actgttgccc acatagaaag 60
atctaaacta attaataat ccctcaccgg caaatcttc agtcaactaag ttagcacat 120
tggtaacaa gttctccaaa ggagagatac agatgagtgc gtatagggtg gacctggctg 180
ctgaggagac acctgcattt gactaagaag agccacgggtg ttagttgaat ggtgtggagt 240
agggtggttc tggggacag tagaaaatcg agaggcatgt gcccgttagt gaactgtatgg 300
aagctacccc aaacgacaga gattgtcagt caggccaaatc cggttcgagt ttgtatggca 360
gcccggacagt gagacagaca cacctactca gttggaggaa ggatgagaac aatggccagc 420
agggattttag agaccctgac aggcgcacagg ccctaacaca cacacctacc acctcaactt 480
acaaagctgc caaagaccaa agacttgcattt tccatttagaa atgacagctg gcttgacccg 540
acagcataat aagcagagtg tactctgatt ggagaacttt aatgtgtttc attcagtatt 600
ataaaaaggac agtattacag attttgcattt acactgtgt tacatgtggg gcagtgtgtc 660
tttaagttagg gtaaaagtact cttaaaaaat gggccctaga tattttttcc tttaactcaa 720
gtctcttact gtttaatga tttttatttt gtttaatatg gggaaaaaag aagcgtaat 780
ggacaatata tatttagaga aagatggta gctgtcagaa aaatatgcaaa atcaaaatca 840
caccaagact gcagcacacc cctgtcagat ggctgtgatc aagaaaataa atgacaatga 900
gtgggtggta agatgtacta aaggaaacaca cacacacaca cacacacaca cacacacaca 960
cacactggag caaccactgt gggaaatcagt atgaatggtc ctcaaaaaacc tgaagataga 1020
gccccggctg gtggcataca cttttattcc cagcaactggg gaggcagagg caggtggatc 1080
tctgagttcc aggcacccctt ggtctatagc acagggttca ggacagccag ggctacacag 1140
aaaaacccctg ctttgattaa accaaacccaa accaaacccaa accaaacccaa accaaacccaa 1200
acccaaacccaa accaaacccag accaaacccaa aacactgaag atagaacttc agtattccat 1260
tccttagat atacccaatg gagactaagt cagcaagaca cctgcacagc catgttcaact 1320
actacactgt tcaccacagc caggctgtgg aaccacccctg agtgcctatc ataaatgaat 1380
ggataggtaa ctccatggatc aaatggactc tgctgtgtac atgcctcaca ttctgtttat 1440
tcattttctt ttatgggtg tccattcagg agtcacatgg tagttctatt ttcaacttcc 1500
tgaagatact acactggtcc ccacagttt cacttttac agcagtgaat aagggttcct 1560
ctatccttac catcatttgc tgtaattttt cttgatgacc ctctttctga cagggatagg 1620
atgttaatatc agtgcgtggaa agtacaactt gttttctaaatg tattttatgg ccccttgcatt 1680
ttcttctttt gaaaactgtc gggttccgtac atctgcgtcag gtattcattt gatgttgcatt 1740
ctttgggttt tgagttctta tgaatttgcattt atgttaatgc cctgcctgtg gttctctccc 1800
attctgttagg ctgcctccctc acccctggcaa ttgttgcctt tggttgcag aaacttttgcatt 1860
cttcatggaa tctcatttgcattt cagttttccctt cctctgtca tagcctgagc taatgcactg 1920
gtttttacag agccctggtc tatgccttta tcctcctctg gcagcttcgg agtttcattt 1980
cttacattta gatcttgcattt ccactttgaa caagtttgg agcagggtga gagatacggaa 2040

tctagttcca ttcttcata tgtgatccta gtttacatag catcggttgg 2100
 ttatTTatt tttaaataat gtgtcataaa aaacgagggt gttgttagcag tgtggatttgc 2160
 tttcttgc ctttgcata caggtcttgc tttgtgtcag tctcatgtat ttttattgtc 2220
 atggctctgt catacagtct gaggtcaggt attgtat accttcagta ttgctccctc 2280
 agactcaggt ttgcttggc caggagtcat cttactcagt gctcttagag ctcccccagc 2340
 atgttagctgc tactattctt agttgataaa tcaggaaact ggggctcaga gagattaact 2400
 gtcttgcact acttctgggg aggtgaaacg tggagacact aaactgtgtt taccctgtac 2460
 tgctccagta gctgtcggt gctgggctac agcaaaacac ctataactata tattactcag 2520
 gaggtggaaa aactcagcct cccttgggg tcccaagctc ccaggtgtcc agtcactgct 2580
 ggaaacctca tggagtctga aaggaagggt tgagggtaca tggggcagcg atgaggagcc 2640
 tggggctggg atctccaaa cacctggata tccagatgcc actgggtcag ggggagttgg 2700
 gaacagagtt gggatgtcca tggacctgtg acaaggccag gcccaggggg aggataactc 2760
 tggcttact aatttgcgaa agtccttagc ttagcagcag ttgtctggg acacagagg 2820
 gccttctgt aagaggctcag gcagtgcgc tctgttaggc aaggtcttct ccatgttccc 2880
 catgggtttt cttgatgaaa gagacagtcc ttggctccaa actgggtttat tgattgttca 2940
 ttgtggaaaa tgggtgcaca ccaccttctc agggtggacc agagatcaa tacctttgc 3000
 agggaggaat atctgggaag ggacgcttac tggctaaacc ctcagggcct ctagatacat 3060
 cattagcatg gagaactctg ttctggcta catgaccaca gcccacattt ccacaagcca 3120
 catgtggaa gtgtggcaca tggcttaggc caggaatctg gttagggagcg tggagccacc 3180
 taccatccca ggtgggtgcc tgggtgccag ggaccctgaa cccgctcaac cttaccaagt 3240
 ttcttggcag ggtccactgt cctacacaga agctggagga ggtgtgaggg ttgtgtctt 3300
 gtggaatgtc ccatgctgtc tggggctcag tttctccacc tgcacccat tgggttgggt 3360
 ataaaaatgt gggataactt attattctct gactcggtcc tgaggaaaaa gcatcggtggc 3420
 agtccagggaa ccacacccctg aggttccctgc actgaaggga ctccctaaatg ctctggagtc 3480
 tctcccttc acagagctgc caaagtctag gttctttga ggataacaga gccatgcttg 3540
 gtaagcagac aacagcattt gtttactcaa ctttcttttgc tcaagtcctt cttcataaac 3600
 aagttgagac accatgctgg cttaggaaag acttctaaag ccagacaact gtcagggaa 3660
 gaagaagaag gggcaagtgg agttgcctg gatgtagccc tcaaagtctc cagagaccag 3720
 ccatgaaggc tcaagtggag ggcaagaccc gcaagccca agcatctggc aggagaggat 3780
 cctggaaacc cctctaccat gacacacatt cttcctgcag gtcacactta ataggccatt 3840
 tcttatttgg atctatcatg gtgttctgtc cgagattaat gaggtgtttat gctgcgaaca 3900
 gaaagttata taaaaacaag tccccccccc ttgtcactgc tgctaaagat gtagcagaaa 3960
 ttgtctcaag tgccttcata atcagaaaca ataaaggctt ctttggattt aagcccttca 4020
 gtttcccttc tccctgtcgc gccttggaca cccatcataaa cttcctggat gctacagctc 4080
 tggcagaga ctccaaagggtc gggagagact gatggtacaa aagcaaaata cttgtttggg 4140
 ggtacaccca ctccctgc tttgtggttc ctgcagtcag tccctgcagac aggccctcag 4200
 tgggtttcc atggcaaca cgcagaggga ggcaatggat ggaataaccc acaccctgg 4260
 tagtttaccc cggccatgtc ctctgcctt catcccttc ctgcctctg ccacggctt 4320
 ctctgcagga atcatatctt catattggcc cacagggtt ctccctcaccc tagctatgt 4380
 gtttacttta gagtgacctt agcagggtc gtggaaatga gttctagaag gtcacagg 4440
 atgttaggaa agaaacgtct tctaactact gaggttacta agttcctgggt gttgtctct 4500
 gccttccct tggtaaaatgtc accttgcactt tagtgcagaa gaaatcagag cccagtcaca 4560
 gagtaaaatgt ggtcctgaag atttccctttt agtgcggcaga atccatgaca tttcaagagc 4620
 cctctttgtt ccttaagtca tttggggttt tatcttctgc ttgtatgtat tgcgtgtgtt 4680
 tatcaaagag tgagatgggtt acataagagg tgctctaaag gacagagagg atttgcaatt 4740
 gtggcatgtg acatccctcag gccttgcctc ggtgccagga ggaactgtatc cagaaaagag 4800
 taagaggtca tttccctggag gctgtacta tagaggagat cttacagtgc attccctcct 4860
 ccaggccctg cctgaggata gacatgtgc gactgcaact gaaacagagg cttggatgg 4920

agatggtaggt tcacagaagg gaggggtggga gatggatgt tgctgggttc tgggtctcat 4980
caccagctcc tgaccacccg gtcagcccat gtgcttattc catagcttc ttttgcata 5040
tttactcagt gtgggtgtttg ttgggaccacca gcagaagcca gtcggcaggct gacagctgt 5100
gatacacagg gcagcatgag ggtcctcagc ctgaagcagt caggctggca gaagagaag 5160
accagcacac attccttcaa ccaactatgt cttgaaaaac aaacatatta ttcacat 5220
attgcattta tgagacagct aaaatgtact cgggtagcat gactccagg 5280
gcaagtgcac tgagtggcag agggacagcc aatgtgag 5340
cacagcttag ctccctgggt ttgggtcaaa ctttgagagt ttgaccacaa gcactttt 5400
tttgacat 5460
ataaaagctt aggcattgact acattaaaat gccttgca 5520
caagaatgtt tctattgact gagaataat gttcaggata aagatccagg aagaaaaagat 5580
cagggataag taaaatacta aactctttg caaagatcat agaccctt 5640
gggttctatt gactgacaag cactgctcag gagttggaa agagtct 5700
tagcctggag actctagtga ggtctagtt tacagacagc aaaaatcacc agttacaaa 5760
ctacattcat ttccagttt ctgatcaggc acaggatgt 5820
aagtccatgt gttttaaaata tctggtttcc 5880
tatacaactc ccacctggag tgacatcttgc ttttcatgtt atattacata cctagacacg 5940
ctcatctcac agacttagga ctttgcattc tgatctccat ttctgtatccc acttccac 6000
ttgccttgat agtgcattt tcttactgc cttgggaca accatgtt 6060
tttgagtg 6120
tgtctggatg catgtaatc tctacacaac aaccctccc cgtcaactcag 6180
attaacacat acatgggt 6240
gctattggc ctttttgact gtccaggaag gcccccaatt gctgagacaa 6300
tcttcattat tgactcta 6360
atcagaggtt ccagtgt 6420
cataaaagca ggaccactgc cccaggaagg tcctggaaac tgatgcagg 6480
gttataaacc 6540
tgcttagggg ctaccaaccc aatccctt 6600
aataaaacagg caggacctgt 6660
agcacctgca ctctcaggat 6720
ccagccttgc gtaccttcac tctcaggta ccccatca 6780
accctgggtc cagccctcag tacctgcgt 6840
cgtctttctt 6900
tttccctca caccctca 6960
tgtggtcctt 7020
ccactatact tcaggggca 7080
tccaacggct 7140
cttcgaaagc 7200
agttgagagt 7260
tctgttaat 7320
ttaggtgtta 7380
atgacagtag 7440
gctttgtgg 7500
ctcttagggccc 7560
aaaggaaaat 7620
tttagttgtg 7680
gtgggtcaac 7740
gtctctcc 7800

caatatggaa gttacagaag tgaaaaaagg cattgcctt gttgggtggtg ttaccagctg 7860
 atttttggtt gtcctgcaag gaggtctggg gactggctgc tctgtctctg tctgtatgag 7920
 tgagggaaagt ctggggagca gattccctaa ctttcagecct ggctgggttc ctgagtgaac 7980
 ccagcctctc tggctctagt actttttcc aaacagaat ctgagtggtg acagggaaaca 8040
 agtaccagcc cattgcctaa gtgccagggt tagtgaggc aggaagctgc catagctggg 8100
 attagtagtt gtattggatg taggaagtcc tttcctggg cagctaattcc ttaatgcctc 8160
 actggagatt ttcaatgaga aatttatccc acggccata tggcccccattt ctggctctc 8220
 caacagccaa gtattttcca tttagaggaga cttcctgtac acttgatgga tgctcatcc 8280
 aaggtgactt gggcagtca gtacagactt gggatgaccc ctgacagccct aacctctccc 8340
 caacaaggcc cctctatgtt tgctatgtaa tgtaatgtca gacattgtca ggagtgtccg 8400
 cagcacagcc tgcccagtgt gagggtctc atagggttcc cactgtctt tctacacagg 8460
 gataacgagg aggttaagctg cagttcccag ttcacttca cagaggaaga gataacccca 8520
 tcccaggtca ttttagccagc agtggaaaga atgaggattt gaactcagg tttccaagtc 8580
 ccattgatac catctcctca caagtccctt gccaccctca cgatgcctt gacacttgcc 8640
 tgccctttat actaaggaga tgcaggtaca aggggtttac ccatgttagca gctgaggcc 8700
 ctggggatag ataccagcag caggcctgtt gtcaccactc taactccagc atccccagtc 8760
 tttgttcctg gagtgtgaaa atccctactt aacaagattt tgcaacagtc cttggctctg 8820
 tgacccatag ctggaaacag gattctcatt gatttggtaa acatgggtggc agccagccaa 8880
 aaagagggtc tgcatacaga agacacgtgt ggcaaggcca cagcagactc tgactaccc 8940
 agcttacaga attacaaggt cataatgtcc tctgctttgg tcacctcatg ttaaggacag 9000
 gcccataatga agatggggca gaagactgaa ggaatggcca accaataact ggcccaactt 9060
 gagacccatc ctacaggcaa gcatcaattt ctgacactac taatgataact ctgttatgct 9120
 tgcagacaga agcctagcat aactatctc cgagaggcacc acccagcaac tgactgaaac 9180
 agaaaaaagat atccacagggc aaacagtggta tggaggtcag ggactattat gggagagctg 9240
 tgggaaggat taaaaaccctt gaaggggata ggaacccac aggaagacca acagagtcaa 9300
 ctaagagacc tttgggagact ctcagagact gagccaccaa ccaaagagca tacacaggcc 9360
 ggtccgagggc acctggcactg ttttttttttccat gtaggtcctc 9420
 caataagccg tagcctgact gcatgtatcca atccccacaa gggctgcata gtctggcctc 9480
 agtgggggag gatgccccta atcctgcaga gacttgatga gtggagagct atccaggggg 9540
 aacccacccctt ctctgagaag ggaatggggta tgggggaggg actctgtgaa gggggacaa 9600
 ggacaaaacaa gaacctcaaa tagtgcaggc cctaaaggct tgcttaaggtag cagtggccca 9660
 gctctgtctt gttccctcagc ccaaggctca gctccacccctt gtttctgtgt ttttctggct 9720
 tttcatggc ctaggacttg ttttttttttccat gtaggtcctc 9780
 tacaagacta gggacattcc ttttttttttccat gctgtatgta ggtggaaagga 9840
 cccaaatcaact ggatttctac ctttcaccaa ctttcacccctt gtttctgtgt 9900
 atttcttcca ctgaaaaatg agactctcaatg atgaacgtcg tgacaaatcat caggcttattt 9960
 aaagaggtgt atctaaccctg aatggcaagc agacagcagg ctttttttttccat gtaggtcctc 10020
 taggaaggac aagaactgct ctttcacccctt gtttctgtgt 10080
 ttttttttttccat gtttctgtgt 10140
 ctttcacccctt gtttctgtgt 10200
 taaaacaaatg ttttttttttccat gtttctgtgt 10260
 gagctgtggg catgaaccac agtggcactg gtttctgtgt 10320
 ggtgagctca gtaaatgtga gtttctgtgt 10380
 gtttttttttccat ttttttttttccat gtttctgtgt 10440
 ctttcacccctt gtttctgtgt 10500
 ttttttttttccat gtttctgtgt 10560
 aactatcagg ctttcacccctt gtttctgtgt 10620
 ttattccagg gggacatcc gtttctgtgt 10680

gacaatggag cctccatcg agtattactt tagctcctca ccgatggcaa tgcaccacct 10740
 ctctacccac ataggagttg ggtctccaca caccccccaca ccccccac caaaacgttt 10800
 tcagttactt tatctggtaa agttcatcg agaatgaagc cagtattaag aacatggaat 10860
 catttggaa cctggatcta gcaatacccc accctagatg gagttgtca gtttgcacct 10920
 cagattataa ttcccccta gcttctatgg tttattctga aaccagggga actcgattcc 10980
 tcccttggaa ccacagacat cctggcttgt gaattcacat gtcatctact gctaatccat 11040
 tggtagtatg tggctcacag agacacacta cagtcatggc caatgtcaag gtaggacaga 11100
 tgtgaatcat tccccccatc ctgctgtttt catgactaac cctccctcagc acagtgcacca 11160
 tgaacctact tttccctcc ttttattttt agaattgtcg gaattttctta ttttgagaaa 11220
 taatagcctt gggcagcatt aaacaaaatc atctagaaaag ctggtttaaa atacagatgg 11280
 ttgagtcagt gaaagagtga ggaatgtcat tattggcccc tcacagaggc tggctcaactc 11340
 cagcagaggt ggttgaagct cttggacacg ggtcagggtgc ataggaaagg tngtctggga 11400
 cactgagaac cacaattgaa caaacagaac tggctgttt tttttttta aatgagttct 11460
 caaaaaaaaatga ctggcttagt taggcaataa cttcgagcca acccaacaga acattctcc 11520
 attgattcat tctggatctt ctttctagac aatactgaac tgaccccttg ttggcagtc 11580
 caagtttgac aacatagggc tttgaacttg gcacaaggc catcaactgtc acccaagcat 11640
 cctgggtgac ctttgggttg gaatatcttg gctaacctta gatattttt ttggagtatc 11700
 ttttagaacat ccagggaaata gggcttgatt ctcatcctgg gaccacaata taagtcaccc 11760
 tagaatccca ggagatcgtg cagagaaaca aggtctctc tcgtgtgcac ctttcttcaa 11820
 agcagtgagt agtactcca ctaaaactgag ttcccatctg agagtccaca ggaggcttg 11880
 gggcaagaag cagagggaaag gcaactgtttg tggtggtaaa gtttgactc taacaaatcc 11940
 gaagacatacg atgacattgt gtcagactaa caacaaccta gactcatgtg gtttctgtt 12000
 agggatcaga ttttattcat caatgacttg tcttagtgc tagagaaagg cttcctactg 12060
 gagtgttaggc tcaataatga cagaagagat agctattcc cctaggact gtgctgcctc 12120
 aagtttggtg gagaaggca gtggggaaacc tagatgtgc ctctggggag ggggtctgaa 12180
 gctggcttca tagaaggtgt gaagtttgc tgaaacatct aaacagaatt atagcttagg 12240
 aaagtggcga ggcaaggcag ggaatgtgtt gcatatgtat atgtacatga atatattatg 12300
 ttatagatac acacacattt gaaaccttgcatt tgcagatgac agaaaatagg ttatttgcc 12360
 tctcttaact gctaagcaca atgacttcca gttccatcca tttcctgaaa tgccacaatt 12420
 tcatttttca ttgtggctga ataaaaattcc attgcagact gggccctact tcattccactc 12480
 ctgagggcag gcatatcccc tggctccatt tcttacctat tggtaagaga agtgcactg 12540
 tcttggtaaa aggcaagcgt gagagaggca ggcactaatt gtgggtttt gtttcttctt 12600
 cctgctatga ctctccattt gtcagaacca aagatcgata aaagccgcca ccatgaaagc 12660
 catcttaatc ccattttat ctcttctgtat tccgttaacc ccgcattctg cattcgctca 12720
 gagtgagccg gagctgaagc tgaaagttgt ggtgattgtc agtgcgtatg gtgtgcgtgc 12780
 tccaaaccaag gccacgcac tgatgcagga tggtaaaaaa gacgcattggc caacctggcc 12840
 ggtaaaactg ggttggctga caccgcgcgg tggtgagcta atgccttact tcggacatta 12900
 ccaacgcac ggtctggtag ccgcacggatt gctggcgaaa aagggtctgcc cgcaacttgc 12960
 tcaggtgcgcg attattgtcg atgtcgacga gctgtacccgt aaaaacaggcg aaggcatttcgc 13020
 cgccgggcgtg gcacctgact gtgcataac cgtacatacc caggcagata cgtccagtc 13080
 cgatccgtta tttaatccctc taaaaactgg cgtttgccaa ctggataacg cgaacgtgac 13140
 tgacgcgcgtc ctcagcaggc caggagggtc aattgtcgac tttaccgggc atcggcaaac 13200
 ggcgttgcgc gaaatggaaac ggggtgtttaa tttccgcaa tcaaaactgt gctttaaacg 13260
 tgagaaacag gacgaaagct gttcattaaac gcaggcatta ccatcgaaac tcaagggtgag 13320
 cgccgcacaaat gtctcattaa ccgggtgcgggt aaggcctcgca tcaatgtcgat cggagatatt 13380
 tctcctgcaaa caagcacagg gaatgcggga gcccgggtgg ggaaggatca ccgattcaca 13440
 ccagtggaaac accttgcataa gtttgcataa cgcgcattt tatttgcatac aacgcacgcc 13500
 agagggttgcg cgcagccgcg ccaccccgat attagatttgc atcaagagacag cgttgcaccc 13560

ccatccacccg caaaaacagg cgtatgggt gacattaccc acttcagtgc ttttatcgc 13620
cgacacatg actaatctgg caaatctcg cgccgcactg gagctcaact ggacgcttcc 13680
cggtcagccg gataacacgc cgccagggtgg tgaactgggt tttgaacgc ggcgtcggct 13740
aagcgataac agccagtgg a ttcagggttc gctggcttc cagactttac agcagatgc 13800
tgataaaaacg ccgctgtcat taaatacgcc gcccggagag gtgaaactga ccctggcagg 13860
atgtgaagag cgaaatgcgc agggcatgtg ttcgttggca gttttacgc aaatcgaa 13920
tgaagcacgc atacccgctt gcagttgtta aggtacccgg ggatcacaac ttgcctctg 13980
aagaggaaga acagaaggat gccacaactc tcctgtggc tacttccag tggttcatac 14040
ttacttctga tggcatttcc ctctagaaag tgctactatc atccacacat ttctacactga 14100
gaccacccaa aggacccctcc caaattctct tcctctgtta gtatgttcca cacctgtac 14160
caccatccca gaattaaaat cctaactgca ctctggcgtg tgacttgcct cagtccttgc 14220
aataagagtt gttggcagtg ccaggcgtgg tggcgcacgc cttaatttcc agcaacttggg 14280
aggcagaggc aggccggattt ctgagttcga gcccagcctg gtctacagag tgagttccag 14340
gacagccagg gctatacaga gaaacccctgt gtcaaaaaac caaaaaaaaaa aaaaaaaagtt 14400
gttggcagag tttgggttat ataccagggt gaggatcaa atgagtggct gaagctgttag 14460
ccagaaggaa cttagaggat agtcataac ttaaaaaagaa atgttagagag tagcagaaac 14520
attgagagag tggcacaca gccactgtgt gaatgtggca gaacacaatc cagccagcta 14580
tacatgcata agtgtatatt ggcgcctatcc tgactgtatga gacacaggaa aacagataga 14640
cggggtagg tggccatggc ctttctgtcc tgcccttcc taagggtcat ctcaagac 14700
tatgtctct taactcttcc attgtactt agcttctaga tatcacctcc agattagtct 14760
ccttgggtac atcgtgatc ctggtgatcc ccagggttcc ctgattccat ctttgtcata 14820
gaggctgcaa cttaaagaggt ctcttaata ctccacaccc tgatgcaaaa aggaagacac 14880
agaagttcac agaggtgaag tgattcatgt aggacatata gtgagcaagc atcagggtcc 14940
ggattatctg actctactct aacttttatg taaatgtgct ttatgcccatt aacactgtca 15000
ttcctgtgt tca gtcgttggc gagactccca agcactctt ggcacaagcc acaattaagg 15060
gactctgaca ctctgcattt attaatttgc attgtggctt ctatgttcc agattcatga 15120
ttgtttactt tccatatacg gctatgaagg gtgtgaggaa attttttgg gacagaattt 15180
gaggcaatcc acctctctca ggaaggctct atctggaaaaa gcttacaact cagggacagt 15240
aactgttaggc ccagtccttg gtgtccaaaa tgggttttat gtttgaatc tgcaagac 15300
tccatgtgt tca aagggtttt aacatggagc ctccctctgg taacactgtta ttggaggctt 15360
ttgagactgg atgctctttt gtcctatgtt ttgctacatc atctgtcaag atatgaccca 15420
ggcatgtac cagctaccac agactatgcc tctccagctt tcatgttctc cccaccatga 15480
tagacttgc tctctaaaa atggaatcaa agcaaaactt tccctgcattt agttttttt 15540
tttctgtta gtgtttggc acagggacaa gaaaacactc aatacagata attagtacca 15600
gagttgaggt tcattgtct agcaagttgg atcaaatttt tagggctttt gaactgattt 15660
ataagagaca tttttagaagag tctgaagctg tgggtacag aagtgtcacc agtttttaag 15720
aatagttaa tacaccatgg gaattgtgaa aatcagaatg ctcacacaaa ggcagacagg 15780
aaaacgtgag catgtggcgt gtgagagggc ataagaagga accttaggggg aaatgagct 15840
gaagccatcc ggctacgtta gggacgtgt gtggctgtc ttggcccatg ccctggcaat 15900
ctgaatgagg ccaaattttt aaggagtggc ctaactcgat tgcagagaa aatatcaaga 15960
cagaccacca ctcaggctat gcccgtttt tgaccgacca gctactctt gcaagctcta 16020
ttgtgaaatt ccagagcaat ttcacatggc tgaagataca tacagttag tgaagtaagg 16080
ggtgtgggtc cctaagtggc tgggtgcataa atctatgttag gtgatgccta agtgcacactt 16140
gataatccaa aatatcagca atgtggaaatg tcttccaagg agacctgttag acacacattt 16200
tagaactttt ctcacatggc taaaatag ctagctagaa atcatttcc gaagaggtta 16260
gtctgagttt cgggtccagg gcaaaatcc agtgcacatggc aggaaggcat tgcagtcagg 16320
agccaaaggt cagctggc tca cattgcata agagtagaga gtcagagttt gatgaaaag 16380
aggatacagg ttataaaaacc tcactgttca ctctcagcaaa tccattttctt cctaaaaaggc 16440

tttacttctt aaagatttta gtcttcaaaa ccagtaccag tagcctggga acaaaagttg 16500
aaacaaatga gcctttgtgg ggcattcac actaaaaaca gggcatcacc taggaggagc 16560
cctgtgtgca gtaggaagtg tggcctctgt gtcaggaatg ctcaggctaa taaggggtcc 16620
tctatcttag gggccctatg aagattcaac aagtagttgt gagaattccc tgtaaatgga 16680
tgctaccaat ttgacatttg tagacctgct attgtgtgct tcttattgg gctctccat 16740
ctcccaactt tccaacccat attccacatt aatcccttcc accaccatgc aacacttaggt 16800
aggagagaag gaaggtaga agagaaagtg ggtatagatc tatttagact aacctctgct 16860
gattaggggc aagtccaaatc gtcattgtca ggataccctcc aaccagcaac cagcaaaacca 16920
gcaaattcaga aacagcaaaa gcagccaaca aggcagcaact aaccaggcagg attggggtcg 16980
gtagcgtggg agcagtcact actggcttc tcatggctt ggcattaata ctctctcaag 17040
aaattccgta atttttcc caccacctga aattccgtaa tttaaatgc aaactatcta 17100
cagctggcaa aaatcacatc tctccttagag cacaagacaa atcatagttt ctggcttattt 17160
gcaatctgaa gcatctcaat atccccacacc tgggattaaa acaaaaacat attcacatca 17220
cataactgtt tttttttcc aattttttat taggtattttt ctttatttac atttcaaatg 17280
ctatccgaa agtcccctat accctccac ctccctgctc ccctacacac ccactccac 17340
ttttgaccc tggagttccc cgtaactggg gcatataaag tttcaagac caaggggcct 17400
ctctcccaag tgatggccga ctaagccatc ttctgctaca tatgcagata gagacacgag 17460
ctctgggggt actagttgt tcatattgtt gttccaccta tagggtcgca gacccttca 17520
gctccttggg tactttgtct agtccttcca ctgggggctc tgtttttat ctaatagatg 17580
actgtgagca tccacttctg tatttgacag gcaactggct agcgtcacat gagccagcta 17640
tatcagggtc ctttcagcaa aaccttgctg gcatgtgcaa tagtgtctgc gtttgggtgg 17700
tgattatggg atggatccac tagttctaga gc 17732

INTERNATIONAL SEARCH REPORT

Int'l Application No
PCT/CA 00/00430

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 A01K67/027 C12N9/16

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 A01K C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the International search (name of data base and, where practical, search terms used)

STRAND, EPO-Internal, WPI Data, PAJ, MEDLINE, BIOSIS, EMBASE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO 97 48812 A (UNITED KINGDOM GOVERNMENT ;YANKE LINDSEY JAY (CA); CHENG KUO JOAN) 24 December 1997 (1997-12-24) page 18, line 27 - line 31 —	1-69
X	WO 99 17610 A (UNIV CALIFORNIA) 15 April 1999 (1999-04-15) page 7, line 12 - line 26 page 10, line 21 -page 11, line 22 —	1-5, 10, 50-52, 62 6-9
Y		—/—

Further documents are listed in the continuation of box C.

Patent family members are listed in annex.

* Special categories of cited documents :

- *A* document defining the general state of the art which is not considered to be of particular relevance
- *E* earlier document but published on or after the International filing date
- *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- *O* document referring to an oral disclosure, use, exhibition or other means
- *P* document published prior to the International filing date but later than the priority date claimed

- *T* later document published after the International filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- *&* document member of the same patent family

Date of the actual completion of the International search

29 September 2000

Date of mailing of the International search report

13/10/2000

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patenttaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Sprinks, M

INTERNATIONAL SEARCH REPORT

Inte ~~lational~~ Application No
PCT/CA 00/00430

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	SELINGER, L. B. (1) ET AL: "The rumen: A unique source of enzymes for enhancing livestock production." ANAEROBE, (1996) VOL. 2, NO. 5, PP. 263-284. , XP000940728 abstract	21,22, 26,28, 29,50
Y	page 267, column 2 page 276, column 1 -page 277, column 1	15,30, 39,40, 45,46, 48,49
X	ZHANG J X ET AL: "Expression of a bacterial endo (1-4)-beta-glucanase gene in mammalian cells and post translational modification of the gene product" BIOCHIMICA ET BIOPHYSICA ACTA-MOLECULAR CELL RESEARCH, (27 JUN 1997) VOL. 1357, NO. 2, PP. 215-224. PUBLISHER: ELSEVIER SCIENCE BV, PO BOX 211, 1000 AE AMSTERDAM, NETHERLANDS. ISSN: 0167-4889., XP000940605 abstract	50,62
Y	WO 99 08539 A (DIVERSA CORP) 25 February 1999 (1999-02-25) Discloses a phytase - relevant for present SEQ ID NO:7 figure 1	11-20, 30,31, 36,37, 39-49, 53-61, 63-69
Y	DATABASE EMBL 'Online' ID: AF062078, 18 May 1998 (1998-05-18) DENNIS ET AL.: "Cloning vector p34S-Cm" XP002148835 Relevant for SEQ ID NO:4 abstract	56,63,65
Y	DATABASE EMBL 'Online' ID: MMSPG, 3 November 1992 (1992-11-03) MIKKELSEN: "M. musculus Psp gene for parotid secretory protein" XP002148836 Relevant for SEQ ID NO:1,7 abstract	1-7, 10-14, 16-69
		-/-

INTERNATIONAL SEARCH REPORT

Int'l. Application No.
PCT/CA 00/00430

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	DATABASE EMBL 'Online! ID: BLCAT3DNA, 26 February 1992 (1992-02-26) LUCKOW: "Plasmid pBLCAT3 gene for beta-lactamase and CAT gene for chloramphenicol acetyltransferase" XP002148837 Relevant for SEQ ID NO:2,6 abstract	47,56, 63,65
Y	DATABASE EMBL 'Online! ID: MMU73190, 4 November 1996 (1996-11-04) LAURSEN ET AL.: "Mus musculus positive acting regulatory region of the parotid secretory protein (PSP) gene" XP002148838 Relevant for SEQ ID NO:7 abstract	1-7, 10-14, 16-69
Y	DATABASE EMBL 'Online! ID: CVCATBLA, 24 January 1992 (1992-01-24) BOSHART ET AL.: "Cloning vector pBLCAT6 encoding chloramphenicol acetyltransferase (CAT) and beta-lactamase (bla) genes" XP002148839 Relevant for SEQ ID NO:3,5 abstract	14,20, 31,36, 44,47, 57-61, 64-69
Y	DATABASE EMBL 'Online! ID: RNRPI5, 12 May 1991 (1991-05-12) LIN ET AL.: "Rat salivary proline-rich protein (RP15) gene" XP002148840 Relevant for SEQ ID NO:3,5 abstract	1-5, 8-14, 16-69
Y	DATABASE EMBL 'Online! ID: ECAPPAA, 5 February 1991 (1991-02-05) DASSA ET AL.: "E. coli periplasmic phosphoanhydride phosphohydrolase (appA) gene" XP002148841 Encodes a phytase - relevant for present SEQ ID NO:3,5,7 abstract	11-20, 30,31, 36,37, 39-49, 53-61, 63-69

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/CA 00/00430

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
WO 9748812	A 24-12-1997	US 5939303 A		17-08-1999
		US 5985605 A		16-11-1999
		AU 716845 B		09-03-2000
		AU 3021697 A		07-01-1998
		BR 9709793 A		11-01-2000
		CA 2257101 A		24-12-1997
		EP 0904385 A		31-03-1999
		NO 985804 A		10-02-1999
		PL 330532 A		24-05-1999
WO 9917610	A 15-04-1999	US 6004944 A		21-12-1999
		AU 9668998 A		27-04-1999
		EP 1030557 A		30-08-2000
WO 9908539	A 25-02-1999	US 5876997 A		02-03-1999
		AU 8904498 A		08-03-1999
		EP 1003379 A		31-05-2000
		US 6110719 A		29-08-2000

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- BLACK BORDERS**
- IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- FADED TEXT OR DRAWING**
- BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- SKEWED/SLANTED IMAGES**
- COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- GRAY SCALE DOCUMENTS**
- LINES OR MARKS ON ORIGINAL DOCUMENT**
- REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- OTHER: _____**

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.